INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



VEHICULOS DE PROPULSION HUMANA

Febrero 1984

350 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 89
Febrero 1984



EDITA Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

Manuel Puigcerver: *Io*; Julio Mangas: *Una comunidad agrícola de la Edad del Hierro en Europa central*; Enrique Cerdá Olmedo: *La hélice del ADN y su lectura*; Luis Mas: *Gravedad cuántica*; Joandomènec Ros: *Alternancia de presas en un ecosistema simple*; Donato Franco Carrera: *Aerodinámica de los vehículos terrestres de propulsión humana*; José Manuel García de la Mora: *Interpretación de las ilusiones ópticas*; Luis Bou: *Juegos de ordenador*; J. Vilardell: *Taller y laboratorio*.

Ciencia y sociedad:

María-Eugenia Armengod y Antonio Pulido

Libros:

Luis Alonso, M. Pilar García Carrasco, Julio A. Alonso y Jesús Mosterín



LA PORTADA

La ilustración de la portada recrea el "Vector Single", un triciclo fuselado que el ciclista pedalea en posición supina, esto es, apoyado en la espalda. El vehículo personifica los avances realizados en los últimos años en la aerodinámica de las bicicletas y otras máquinas a pedal (véase "Aerodinámica de los vehículos terrestres de propulsión humana", por Albert C. Gross, Chester R. Kyle y Douglas J. Malewicki, en este mismo número). El "Vector Single" fue diseñado y construido por un equipo encabezado por Allan A. Voigt, presidente de la Versatron Research, Inc. Según Voigt, este triciclo alcanza una velocidad de casi 100 kilómetros por hora en suelo llano, con una potencia de 1 caballo de vapor del ciclista. Su compañero, el "Vector Tandem", llevado por dos pilotos situados espalda contra espalda, tiene el récord de velocidad de 101,24 km/h, alcanzado con una potencia ligeramente superior a 1 CV por cada piloto. Consiguió una media de 81,2 km/hora sobre 64 kilómetros en una autopista interestatal.

Suscripciones:
Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 37

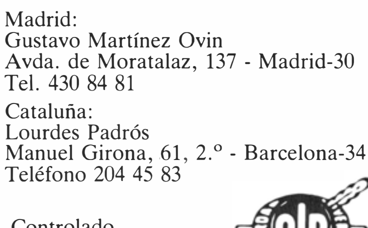
Condiciones de suscripción:
España:
Un año (12 números): 3.850 pesetas
Extranjero:
Un año (12 números): 52 U.S. $
Ejemplar atrasado ordinario:
350 pesetas
Ejemplar atrasado extraordinario:
450 pesetas

Distribución para España
Distribuciones de Enlace, S. A.
Ausias March, 49 - Barcelona-10

Distribución para los restantes países:
Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137 - Madrid-30
Tel. 430 84 81
Cataluña:
Lourdes Padrós
Manuel Girona, 61, 2.° - Barcelona-34
Teléfono 204 45 83

Controlado
por O.J.D.



PROCEDENCIA DE LAS ILUSTRACIONES

Dibujo de la portada de Ted Lodigensky

Página	Fuente
6	A. Guerra y G. Pérez-Gandaras
8	Recreación a partir de W. Patten ("Embryology of Patella", *Arbet. Zool. Inst. Univers. Wien*, 6, 1886), *A*; Shearer (*B*) y Werner (*C* y *D*)
10	Ricardo Génova, A. Guerra y G. Pérez-Gandaras
11-14	C. Chun. "Die Cephalopoden". *Wiss. Ergebn. dt. Tiefsee-Exped. "Valdivia"*. I: 18; 1910
17-24	National Aeronautics and Space Administration, Servicio Geológico de los Estados Unidos
25-27	Ian Worpole
28	John T. Trauger, Instituto de Tecnología de California
31	Caroline Quillian Stubbs
32-36	Alan D. Iselin
43	Nelson Max, Lawrence Livermore National Laboratory
44-48	Irving Geis

Página	Fuente
49	Richard E. Dickerson, Instituto de Biología Molecular, Universidad de California en Los Angeles
50	Irving Geis
51	Richard E. Dickerson, Instituto de Biología Molecular, Universidad de California en Los Angeles
52	Irving Geis
53	Richard E. Dickerson, Instituto de Biología Molecular, Universidad de California en Los Angeles
54-56	Irving Geis
59-69	George V. Kelvin
73-80	Tom Prentiss
83	Bruce A. Sargeant
84-85	Walken Graphics
86-89	Albert E. Miller
90	Walken Graphics
92-98	Gabor Kiss
101-106	Ilil Arbel
109-110	Frode Wissing
111-114	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Nápoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Los autores

ANGEL GUERRA y GERMAN PEREZ-GANDARAS ("Larvas de cefalópodos") trabajan en el Instituto de Investigaciones Pesqueras de Vigo del Consejo Superior de Investigaciones Científicas (CSIC). Ambos se dedican al estudio de los cefalópodos desde hace nueve años. Han realizado numerosas campañas de investigación nacionales e internacionales, y han publicado más de setenta trabajos sobre este tema, algunos de ellos premiados por organismos oficiales. Guerra es doctor en ciencias biológicas por la Universidad de Barcelona, colaborador científico del CSIC y jefe del equipo de biología pesquera. Pérez-Gandaras es doctor en ciencias biológicas por la Universidad Complutense de Madrid con un trabajo sobre taxonomía de cefalópodos a partir de sus mandíbulas. El primero se interesa en la actualidad en temas de ecología y citotaxonomía de los cefalópodos y el segundo en problemas de taxonomía numérica mediante ordenador y en modelos de crecimiento.

TORRENCE V. JOHNSON y LAURENCE A. SODERBLOM ("Io") son, respectivamente, astrónomo y geofísico; formaron parte de los grupos de la NASA que obtuvieron las primeras imágenes a corta distancia de muchos cuerpos del sistema solar. Johnson se doctoró en ciencias planetarias por el Instituto de Tecnología de California. Antes de acceder a su actual cargo de investigador para el Jet Propulsion Laboratory del Instituto de Tecnología de California perteneció al Laboratorio de Astronomía Planetaria de la NASA y al Consejo Nacional de Investigación de los Estados Unidos. Colaboró en el envío de sondas a los planetas exteriores y con el grupo de obtención de imágenes procedentes de los Voyager. Formado en el Instituto de Minería y Tecnología de Nuevo México, Soderblom se doctoró también en ciencias planetarias por el Instituto de Tecnología de California. Desarrolla su trabajo en el Servicio Geológico de los Estados Unidos.

PETER S. WELLS ("Una comunidad agrícola de la Edad del Hierro en Europa central") es profesor adjunto de antropología de la Universidad de Harvard y conservador de arqueología europea del Museo Peabody. Se ha movido siempre en el círculo de la Universidad de Harvard, donde estudió toda la carrera, doctorado inclusive, que recibió en 1976. Asistió también, para completar su preparación, a la Universidad de Tübingen. Su interés científico se centra en la economía y los cambios sociales que se produjeron en las edades del Bronce y del Hierro. Las excavaciones de Baviera que describe en este artículo constituyen la base de irradiación hacia otros asentamientos prehistóricos de la zona que le permitan reconstruir la sociedad centroeuropea del último milenio a. de C.

RICHARD E. DICKERSON ("La hélice del ADN y su lectura") es profesor de bioquímica y geofísica y director del Instituto de Biología Molecular de la Universidad de California en Los Angeles (UCLA). Estudió en el Instituto Carnegie de Tecnología y se doctoró en química física por la Universidad de Minnesota. Tras un año en el Laboratorio Cavendish de la Universidad de Cambridge regresó a los Estados Unidos, contratado por la Universidad de Illinois. Cuatro años después se trasladó al Instituto de Tecnología de California (Cal Tech). En sus 18 años de estancia en Cal Tech sus principales temas de investigación fueron la estructura cristalina del citocromo *c*, una proteína transportadora de electrones, y la evolución molecular. Desde su traslado a UCLA en 1981 su interés se centra en la estructura de fragmentos cortos de ADN y de los complejos que éstos forman con drogas antitumorales.

BRYCE S. DEWITT ("Gravedad cuántica") enseña física en la Universidad de Texas en Austin, cuyo centro de relatividad dirige. Se formó en Harvard, por la que es doctor en física. Tras una breve estancia en varias instituciones obtuvo una plaza de investigador físico en el Laboratorio de Radiación de la Universidad de California en Berkeley. De allí se trasladó al departamento de física de la Universidad de North Carolina en Chapel Hill. Estuvo 15 años en North Carolina. En 1972 lo contrató la Universidad de Texas. Ha sido profesor visitante en muchos centros: École d'Été de Physique Théorique de Haute Savoie, Francia, Universidad de Osaka, Universidad de Stanford, Universidad de Oxford e Instituto de Física Teórica de California en Santa Bárbara.

ARTHUR T. BERGERUD ("Alternancia de presas en un ecosistema simple") tiene plaza de investigador en la Universidad de Victoria, Columbia Británica (Canadá). Se graduó por el Oregon State College en 1956 y se licenció en biología por la Universidad de Wisconsin cinco años después. Se trasladó seguidamente a Terranova para incorporarse al Departamento provincial de la Naturaleza. En 1966 se matriculó en la Universidad de Columbia Británica para ampliar estudios. Se recibió de doctor en biología por esa misma universidad en 1968; su tesis versó sobre la población de caribúes en Terranova.

ALBERT C. GROSS, CHESTER R. KYLE y DOUGLAS J. MALEWICKI ("Aerodinámica de los vehículos terrestres de propulsión humana") comparten su interés en el tema de su artículo, al que se acercan desde perspectivas profesionales completamente diferentes. Gross, consejero de dirección y periodista, obtuvo dos licenciaturas: una en historia de América por el Columbia College (1969) y otra en psicología por la Universidad estatal de San Diego (1975). Prosiguió sus estudios de psicología experimental en la Universidad de California en San Diego. Kyle es profesor de ingeniería mecánica de la Universidad estatal de California en Long Beach. Licenciado por la Universidad de Arizona, se doctoró en ingeniería por la de California en Los Angeles. Bicicletas fuseladas diseñadas por él se han utilizado para establecer cinco récords mundiales de velocidad. Malewicki es director de Aerovisions, Inc. Se licenció en ingeniería aeronáutica y astronáutica por Standford. Es consejero de proyectos aeroespaciales, incluidos los relacionados con la lanzadera espacial. Trabajó en el análisis de esfuerzos en submarinos biplazas y análisis de actuaciones de globos de aire caliente diseñados para establecer récords de altura.

DONALD D. HOFFMAN ("Interpretación de las ilusiones ópticas") es profesor adjunto de inteligencia artificial y ordenadores de la Universidad de California en Irvine. Se licenció, en 1978, por la Universidad de California en Los Angeles y se recibió de doctor en psicología por el Instituto de Tecnología de Massachusetts en 1983.

Hace...

José M.ª López Piñero

... trescientos años

Pierre Bayle expuso, en su célebre revista *Nouvelles de la Republique des Lettres*, las ideas de Gómez Pereira, considerándolo el primer autor que había afirmado "que les bêtes sont des machines" y planteando el problema de su posible influencia sobre Descartes. Esta exposición la amplió el propio Bayle doce años más tarde en un artículo consagrado a Pereira del *Dictionnaire Historique et Critique*. El importante peso que tuvo Bayle, a través de sus dos publicaciones citadas, en las corrientes racionalistas y críticas propias de la Ilustración, explica que ambas exposiciones contribuyeran decisivamente a la difusión del pensamiento del autor español entre los estudiosos de temas antropológicos y biológicos. Una función parecida había desempeñado en los ambientes médicos la cita del inglés Thomas Willis en su tratado *De anima brutorum* (1672) –título fundamental en la historia de la neurofisiología y de la anatomía comparada del sistema nervioso–, en la que también presentó a Gómez Pereira como primer defensor de la teoría según la cual los animales carecen de sensación.

Gómez Pereira nació en 1500 en la localidad castellana de Medina del Campo. Era hijo de un comerciante llamado Antonio Pereira, quizá de origen judeoconverso, y de Margarita de Medina, también de una familia de mercaderes. Estudió artes y medicina en la Universidad de Salamanca, donde fue discípulo de Juan Martínez Silíceo, del que aprendió filosofía natural de orientación nominalista, y asimiló el saber médico tradicional de acuerdo con los supuestos del galenismo bajomedieval anterior a la corriente humanista. Al terminar sus estudios volvió a Medina, donde permaneció el resto de su vida dedicado al ejercidio de su profesión y también a atender algunos negocios familiares. Llegó a tener gran prestigio como clínico, como lo demuestra el hecho de que fuera llamado a la corte para asistir al príncipe Don Carlos, aunque no llegó a desempeñar ningún puesto en la casa real. Fue, por el contrario, el prototipo de médico e intelectual plenamente integrado en el estamento de los burgueses de Medina, durante los años de florecimiento comercial de la ciudad. Ello influyó en el carácter de su obra científica, que no puede equipararse a la procedente de la cultura universitaria de esta época.

Gómez Pereira publicó únicamente dos libros, ambos resultado, al parecer, de una preparación a lo largo de varias décadas. El primero, que tituló *Antoniana Margarita* (1554) en honor de sus padres, es una de las obras más famosas de la historia del pensamiento español. A pesar de ello, no había sido debidamente analizada hasta las recientes investigaciones de José Bernia. Su tema central es la tesis de la insensibilidad de los animales ("bruta sensu carent"), que el medinense fue el primero en defender. Con un razonamiento que Laín Entralgo ha comparado al famoso de la estatua de Condillac, Gómez Pereira afirma que si los animales sintiesen, tendrían capacidad de juzgar y no se diferenciarían del hombre: "Si los brutos ejerciesen los actos de los sentidos exteriores como el hombre, al ver un perro o un caballo a su dueño, concebiría lo mismo que concibe un hombre cuando ve a su señor, y afirmaría mentalmente que aquél era su dueño... luego habría que concederle la distinción, que es la más importante entre las operaciones racionales... Si los brutos sienten como nosotros y juzgan, componen y distinguen... ¿Dónde queda la diferencia específica del hombre? ¿Por qué le llamamos animal racional?". La conclusión es que los animales son autómatas –aunque Gómez Pereira no utilice esta palabra–, que "se determinan a obrar y a moverse mediante ciertas cualidades transmitidas por los objetos extrínsecos a los órganos de los sentidos". Esta tesis, que le lleva a estudiar el comportamiento animal con un detalle inédito, supone una determinada filosofía natural. Bernia ha desmentido el atomismo y también el mecanicismo que Menéndez Pelayo y otros autores habían atribuido a Gómez Pereira. Su física continúa siendo la tradicional, aunque sustituya la materia prima aristotélica por los cuatro elementos considerados en su conjunto, de donde procederían las cualidades primarias. Básicamente, es la filosofía natural nominalista que había aprendido de un seguidor de los "calculatores" como Martínez Silíceo. También la anatomía y la fisiología del sistema nervioso en la que se apoya es la del galenismo bajomedieval que había estudiado en Salamanca. Su rebelión frente a la autoridad de Aristóteles y Galeno es muy distinta a la surgida en el seno del movimiento humanista. Afirma que Erasmo de Rotterdam y los demás autores de su mentalidad se limitan a una mera discusión terminológica, en lugar de interesarse por el discernimiento de la realidad. Considera que ha conseguido romper la "fascinación" de los clásicos, actitud que puede servir de ejemplo a los "que, como yo, se esfuercen en investigar la verdad". De todas formas, recomienda cautela en la crítica de las doctrinas antiguas, que reserva para la edad madura tras una vida dedicada al estudio.

La *Antoniana Margarita* motivó dos críticas inmediatas. Miguel de Palacios, catedrático de teología en Salamanca, atacó a Pereira desde la ortodoxia académica en sus *Objectiones adversus nonnulla ex multiplicibus paradoxis Antonianae Margaritae* (1555). Francisco de Sosa, doctor avecindado en la propia Medina, defendió "que los brutos sienten y por sí mueven", en su *Endecálogo contra Antoniana Margarita* (1561), sátira destinada a un público burgués. Más tarde, la tesis de Gómez Pereira acerca de la insensibilidad de los animales fue expuesta y criticada por el médico Francisco Valles en *De sacra Philosophia* (1587) y por el filósofo y teólogo Francisco Suárez en *De anima* (1621), obras ambas de gran difusión en toda Europa.

En 1641, el propio Descartes, en una carta a una personalidad tan significativa como Marin Mersenne, afirmaba que no conocía el libro del médico de Medina y que no tenía interés en leerlo. A partir de esta época, sin embargo, se relacionó de modo habitual la *Antoniana Margarita* con la obra de Descartes, aunque solamente sus adversarios lo hicieron para cuestionar su originalidad, como ya a finales del siglo XVII hizo el obispo Pierre Daniel Huet en su *Censura philosophiae cartesianae* (1689).

A mediados del siglo XVII, el jesuita Rodrigo de Arriaga y el médico Isaac Cardoso, dos españoles simpatizantes del atomismo que desarrollaron la mayor parte de su labor intelectual y científica en otros países europeos, se refirieron también a Gómez Pereira como pensador antiaristotélico. Durante la centuria siguiente, tras los comentarios de Bayle y Willis, la obra del médico de Medina se convirtió, en manos de Benito Jerónimo Feijóo, Carlo Denina, Juan Pablo Forner y otros autores, en uno de los tópicos de los apologistas de la ciencia española. A ello respondió la reimpresión de sus dos libros en 1749. El enfoque citado perduró en la segunda fase de la polémica sobre la ciencia española, durante la cual Marcelino Menédez Pelayo publicó su estudio sobre la *Antoniana Margarita*.

El segundo libro de Gómez Pereira, *Nova veraque medicina experimentis*



ANTONIANA MAR-
GARITA, OPVS NEMPE PHY-
SICIS, MEDICIS, AC THEOLOGIS
NON MINVS VTILE, QVAM NECESSA-
RIVM, PER GOMETIVM PEREIRAM, ME-
DICVM METHINAE DVELLI, QVAE
Hiſpanorum lingua
appellatur, nunc primùm
in lucem æditum.

Anno M.D.LIIII. decima quarta
die Menſis Auguſti.

Portada de la primera edición de **Antoniana Margarita** *(1554), de Gómez Pereira. Es una de las obras más famosas del pensamiento científico español, aunque el número y la calidad de estudios a ella dedicados no están a la altura de su celebridad.*

et evidentibus rationibus comprobata (1558), está fundamentalmente dedicado al estudio de las fiebres. A pesar de su arrogante título, se enfrenta con la piretología tradicional con un método puramente escolástico. Lo mismo que en la *Antoniana Margarita*, su autor utiliza ampliamente la argumentación "ad absurdum". Los supuestos básicos son los que ya hemos expuesto. Critica a los seguidores del movimiento humanista por haber convertido la medicina en mera destreza filológica y se sitúa en un galenismo dominado por autores bajomedievales como el avicenista Gentile da Foligno.

Gómez Pereira afirma que el calor febril no se distingue del natural más que por su cantidad. Define la fiebre como un "calor inmoderado... engendrado en todo el cuerpo por la naturaleza, *que se esfuerza en producirlo*". Enuncia así un elemento básico de la concepción moderna de la fiebre como proceso reactivo. Ofrece una nueva interpretación de su producción, basada en la taquicardia como fenómeno primario de la respuesta del organismo.



Larvas de cefalópodos

El desarrollo embrionario de los cefalópodos difiere del que caracteriza al resto de los moluscos. Las larvas de aquéllos presentan un gran primitivismo, con algunas curiosas adaptaciones morfológicas y funcionales a su género de vida

Angel Guerra y Germán Pérez-Gandaras

Cualquier alumno de biología o persona aficionada a la historia natural, que observase por primera vez un cefalópodo recién nacido, no dudaría en identificarlo como tal. Muéstreseles a esos mismos sujetos una trocófora: difícilmente le dirán que se trata de la larva de un molusco, de una lapa por ejemplo. Y, sin embargo, tanto los cefalópodos como las lapas son moluscos. En este grupo de organismos hay dos tipos de huevos, unos con vitelo escaso y repartido por igual, huevos isolecíticos, y otros en los que el vitelo, abundante, se acumula en el polo vegetativo de la célula rodeando al polo animal, huevos telolecíticos. Estos últimos son propios de los cefalópodos; los primeros, del resto de los moluscos.

Se ha observado que en los huevos isolecíticos la segmentación, o división celular en blastómeros, afecta a toda la célula; se asiste a una segmentación total. En los huevos telolecíticos, por contra, el citoplasma no puede integrar todo el vitelo en el proceso de la segmentación; se forma un disco embrionario, que flota sobre el vitelo, siendo la segmentación de tipo parcial. Los primeros blastómeros formados en la segmentación de tipo total sufren una rotación de cuarenta y cinco grados hacia la derecha y hacia la izquierda, alternativamente; por cuya razón se habla de una segmentación espiral. La segmentación del huevo de los cefalópodos es de tipo discoidal, limitada al disco embrionario

La larva trocófora se forma después de una segmentación completa y espiral de huevos isolecíticos. Una vez constituida, la trocófora presenta forma bicónica, con una banda de cilios situada por encima de la boca y otra banda ciliada rodeando el ano. El tubo digestivo, en *U*, está tapizado por endodermo ciliado. La trocófora posee nefridios, o riñones larvarios, un anillo nervioso principal que ciñe a la larva por debajo de la primera banda ciliada y varios nervios radiales derivados de un glanglio apical, así como una zona sensorial asimilable a un ojo. Muestra un régimen de vida planctónico. Se supone que esta forma larvaria constituye un tipo ancestral que refleja las relaciones, aún desconocidas, de un gran número de *phyla* (anélidos, platelmintos, artrópodos, moluscos y otros).

En algunos grupos de moluscos los huevos eclosionan en un estado larvario más tardío y evolucionado que las trocóforas. Estado que recibe el nombre de veliger. La transformación de una trocófora en una veliger requiere cambios de diferenciación y de organización considerables. La veliger está perfectamente adaptada a la vida planctónica; nada con entera libertad. A medida que se va desarrollando, su pie crece mucho mientras que el velo mengua; en las postreras fases del desarrollo, la veliger puede nadar mediante los cilios y reptar con ayuda del pie.

1. LARVA DE CALAMAR (*Loligo vulgaris*) viva de 3,5 milímetros de longitud total. Procede de una puesta de la ría de Vigo recogida a treinta metros de profundidad. Los huevos hicieron eclosión a bordo del barco en que se efectuó la captura.

Algunas de las estructuras que se forman durante el desarrollo embrionario reflejan relaciones con otros *phyla* de organismos inferiores; la mayoría, empero, debe considerarse especializaciones estructurales adaptativas, reflejo de la solución que estas larvas han dado al desafío de una supervivencia en determinados ambientes. Estos caracteres, que se conocen con el nombre de cenogenéticos, son fruto de la evolución. Cuando se los estudia en las larvas trocófora y veliger, uno tiene la impresión de hallarse ante un panorama evolutivo diferente y muy alejado del que ofrecen los adultos. Por otro lado, la relativa homogeneidad estructural de estas larvas en distintos *phyla* de organismos revela que son tipos funcionales de gran éxito evolutivo.

En los moluscos, a excepción de los cefalópodos, existe una auténtica metamorfosis, lo que significa que atraviesan un desarrollo indirecto. Aunque algunas especies de caracoles pulmonados tienen una descendencia que ya está completamente formada al nacer, ello se debe a que los huevos se desarrollan en la parte inferior del oviducto, pasando antes por el estado de veliger; no poseen, pues, ningún desarrollo directo propiamente dicho, como sí ocurre en los cefalópodos. Las fases embrionarias de estos últimos son de tal naturaleza que resulta imposible determinar su parentesco con el resto de los moluscos a partir de las mismas; más fácil se hace descubrir la afinidad entre cefalópodos y anélidos, cuyas larvas trocóforas son parecidas, que entre cefalópodos y demás clases de moluscos.

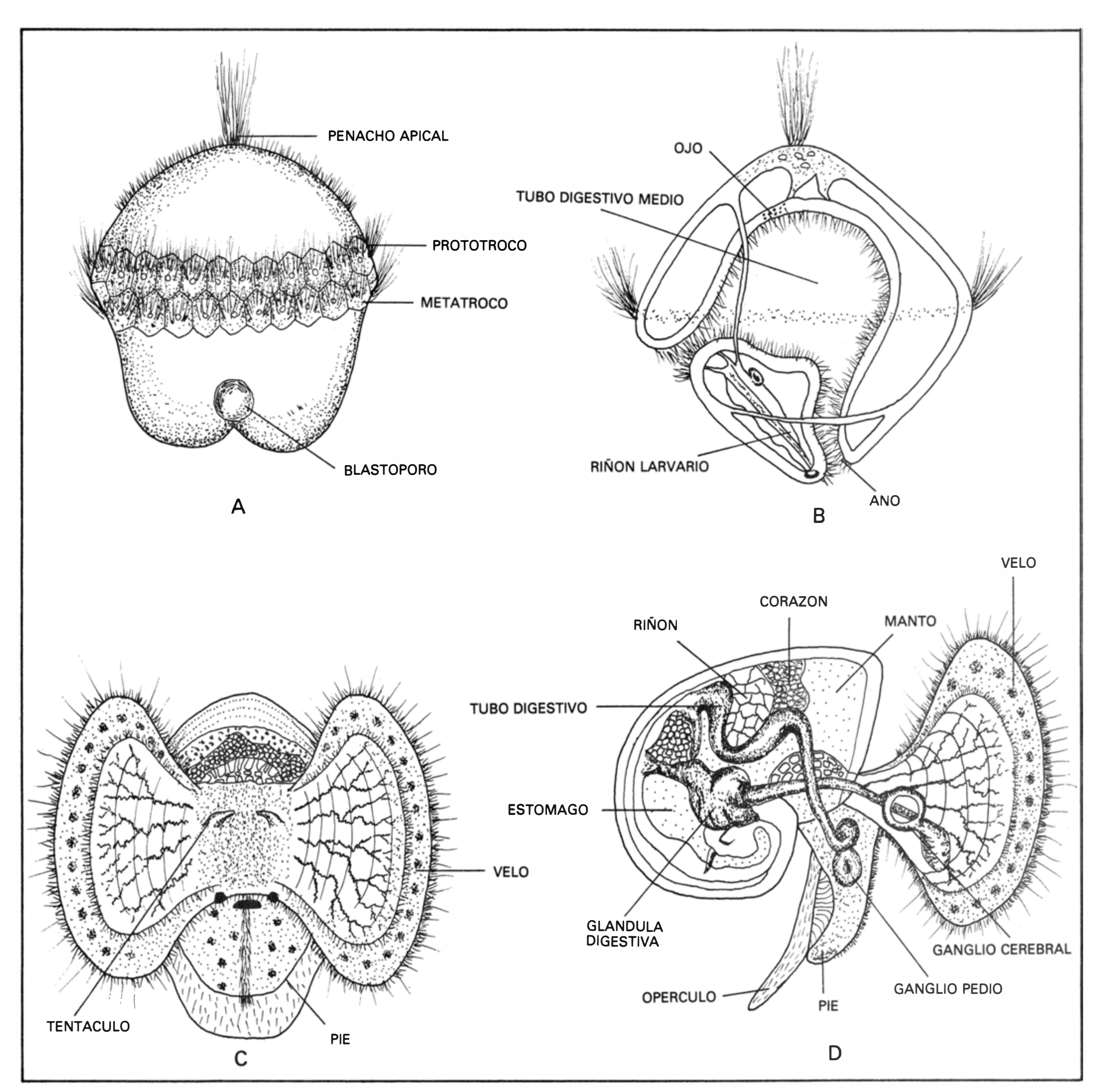
La duración del período embrionario depende fundamentalmente del tamaño del huevo y de la temperatura del agua. Una vez finalizado el desarrollo, la larva perfora la envoltura del huevo o corion, presionando la pared con la parte posterior del cuerpo donde hay una zona glandular que se rompe dejando libres un complejo de enzimas. El órgano de Hoyle, denominación técnica de esta zona glandular, tiene forma de ancla en los cefalópodos decápodos, y de franja recta en los octópodos. Por ello, la perforación adquiere forma triangular en los primeros y recta en los segundos. Cuando además del corion existen otras envolturas, el animal las rasga con la parte posterior del cuerpo, para lo que se ayuda con los brazos que presionan desde el lado opuesto. Los huevos de los octópodos carecen de envolturas externas, pero la fisura abierta por el órgano de Hoyle suele ser tan estrecha que únicamente permite el paso de la parte posterior del cuerpo del animal. Si el pulpo tiene los brazos largos se ayuda con ellos para agrandar la hendidura; si, por el contrario, son cortos utiliza los pequeños manojos de quetas distribuidos por

todo el cuerpo, el órgano de Kölliker, que funcionan cual lima movida por las contracciones del cefalópodo.

Algunas especies tienen estadios juveniles en que presentan ciertas estructuras adaptativas al medio. Pero, en general, sus larvas apenas tienen especializaciones propias y, en este sentido, son mucho menos evolucionadas que las larvas del resto del *phylum*.

El primitivismo adaptativo de los estadios juveniles de los cefalópodos induce a pensar que los huevos telolecíticos surgieron antes que los isolecíticos. En efecto, es evidente que una larva trocófora o veliger que sale del huevo perfectamente adaptada a la vida planctónica, capacidad que adquirió a través de un proceso de selección lento y complejo, depende menos para sobrevivir de la existencia de una despensa alimentaria (vitelo) que otras larvas que, más semejantes a los adultos, presentan una menor adaptación típicamente larvaria. En este sentido se puede afirmar que cuando en el ciclo vital de una especie hay estados larvarios muy especializados, dicha especie está más evolucionada en lo concerniente a esa fase que otras de la misma clase cuyo desarrollo sea directo. En los gusanos nemátodos encontramos ejemplos de ambos tipos de situación: los nemátodos libres gozan de un desarrollo directo, mientras que los nemátodos parásitos del hombre presentan distintos tipos de larvas especializadas para vivir en la sangre, los pulmones, intestino o vasos linfáticos. Los nemátodos de ciclo biológico más complejo han requerido mayor tiempo de evolución que los del ciclo sencillo, más primitivos, por ende.

De ahí que creamos que los cefalópodos conservan huevos y desarrollo em-



2. LARVAS TROCOFORA Y VELIGER de moluscos. Arriba, a la izquierda, se esquematiza la trocófora de *Patella*; a su derecha, una sección de la misma que pone en evidencia sus diferentes partes. Los dos esquemas inferiores corresponden a la vista frontal y lateral de una larva veliger de *Crepidula*, en las que se detalla su organización. Estos tipos de larvas se dan en distintos *phyla* de organismos, lo que indica que son tipos funcionales de gran éxito evolutivo. Su constitución, muy diferente de la que muestran los adultos, se explica por las importantes adaptaciones larvarias al medio que poseen.

brionario con caracteres primitivos. También nos parece verosímil que tuviesen este tipo de huevos y de desarrollo los antecesores de los organismos celomados (moluscos, anélidos y artrópodos). El proceso evolutivo que probablemente condujo a la aparición de huevos con escaso vitelo debió ser gradual. La selección de las larvas mejor adaptadas al medio por la adquisición progresiva de estructuras especializadas para la flotación, natación y alimentación, iría obviando la necesidad de un vitelo abundante en los huevos. La pérdida progresiva de vitelo supondría un considerable ahorro energético, y la aparición de estos estados larvarios podría verse como beneficiosa para aumentar la dispersión de la especie. La energía que se ahorraría en el proceso mencionado puede canalizarse hacia una mayor fecundidad; se comprueba, efectivamente, que las puestas de la mayoría de los moluscos no cefalópodos son de millones de huevos.

El proceso evolutivo en la fase del desarrollo larvario es independiente del seguido por los adultos. Se explica así que, a pesar de sus larvas primitivas, los cefalópodos actuales sean los más evolucionados de cuantos invertebrados existen. En cualquier nivel en que se realice el análisis, se hallará un gran parecido entre ellos y los peces modernos. Aunque poseen la organización y la fisiología de un molusco, han ido incorporando sistemas conformados de manera parecida a la de los peces. Entre estos sistemas destacan el nervioso central y los órganos de los sentidos. Asimismo, la convergencia adaptativa entre cefalópodos y peces es muy acusada en lo referente al comportamiento. Ambos grupos de organismos, filogenéticamente muy distantes, hubieron de evolucionar en una amplia zona adaptativa común, donde la interacción fue muy estrecha; allí, la competencia por el alimento y por el espacio y las relaciones depredador-presa constituirían las principales fuentes de presión selectiva. La capacidad evolutiva de los cefalópodos ha sido, pues, muy notable, conquistando todos los ambientes marinos. El paso de un ecosistema maduro como el bentónico, donde vivieron los primeros cefalópodos, a otro que lo era menos, el pelágico, va asociado, en terminos evolutivos, a precocidad en la reproducción y acortamiento de la vida, que es lo que ocurre en los cefalópodos.

La información contenida en el conjunto de genes o genoma de los cefalópodos conjuga perfectamente larvas primitivas con adultos evolucionados. Diríase, en este sentido, que el bloqueo de unas determinadas vías potenciales de evolución favoreciese el desarrollo de otras, como si en el genoma de las especies existiese una tasa máxima de variación, de manera que, al encarrilarse la evolución por una determinada dirección adaptativa, el resto de las vías potenciales de variación quedasen bloqueadas. De acuerdo con esta hipótesis, cuando determinados organismos entran en la vía donde se complica el ciclo larvario y se producen larvas especializadas, queda mermada su posibilidad para avanzar por otras vías de evolución que, en algunos casos, coincidirían con la adquisición de una mayor libertad por el desarrollo del sistema nervioso, y, en otros, con la posibilidad de una gran radiación adaptativa. Los cefalópodos y los peces, con la relativa simplicidad de sus etapas larvarias y el buen desarrollo de su sistema nervioso respectivo, constituyen dos buenos ejemplos de lo expuesto para el ámbito marino, como los saurópsidos lo son para el terrestre.

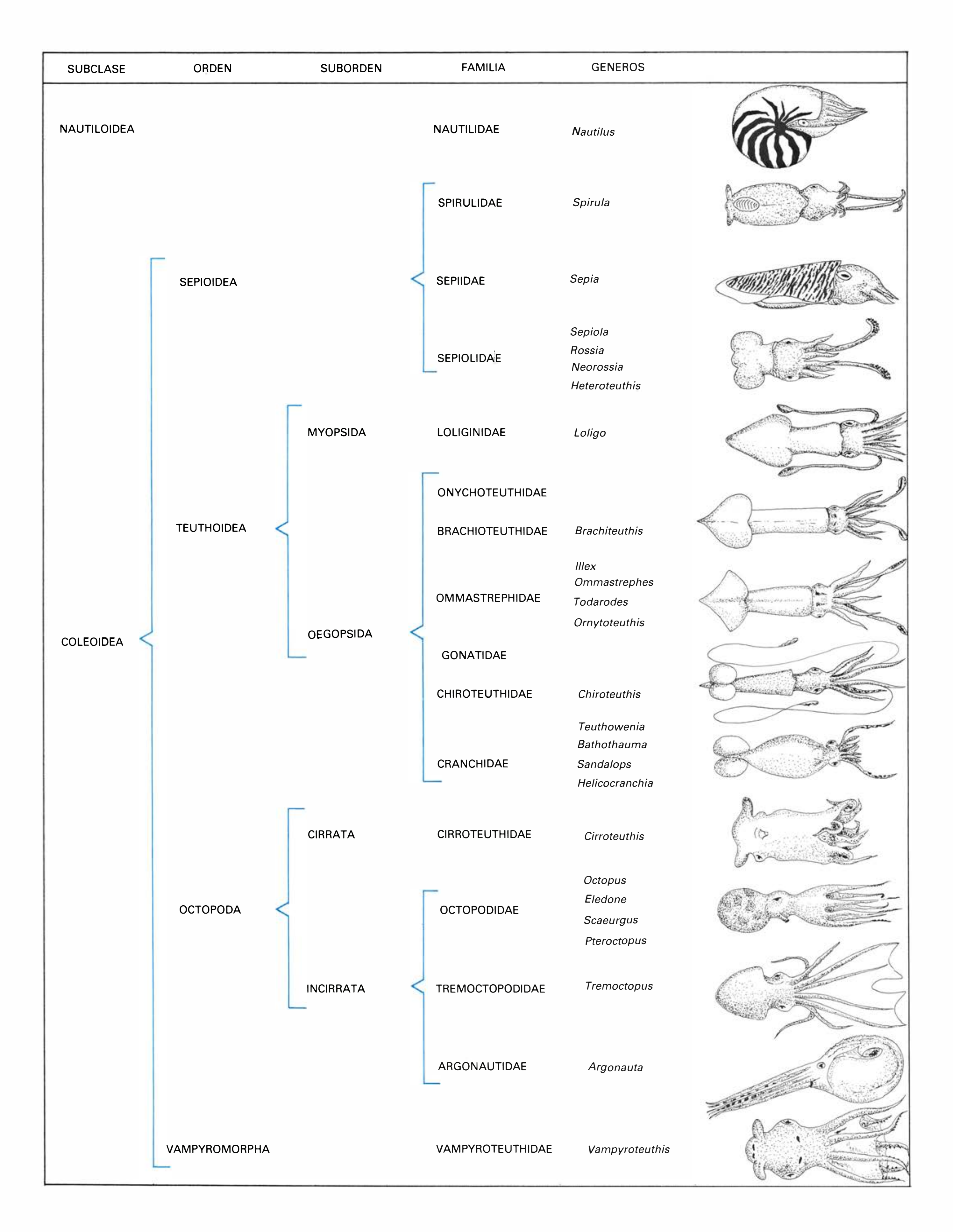
El parecido de las larvas a los adultos, de que hablábamos antes, varía de unas especies a otras. Las más cercanas a los adultos son propias de las especies en que el tamaño de los huevos, con respecto al del animal totalmente formado, es grande; se cumple esta relación en las sepias, en la mayoría de los sepiólidos y en varias especies de pulpos con larvas bentónicas. Por contra, descendientes menos parecidos son típicos de las especies de calamares, potas y algunos pulpos de huevos pequeños y larvas planctónicas. En este último modo de descendencia, las diferencias con los adultos afectan a estructuras que son adaptaciones a la vida planctónica. Por su relativa disparidad de las formas adultas, por las transformaciones que experimentan y por su género de vida, este tipo de descendientes se han denominado "larvas planctónicas"; se reserva el apelativo de "juveniles bentónicos" a los descendientes nacidos de huevos grandes, evitando aplicarles el término de "larva".

Entre los teutólogos, naturalistas que se dedican al estudio de los cefalópodos, persiste la vieja controversia sobre si se puede o no hablar de auténticas larvas a propósito de los cefalópodos. En sentido lato, una larva es un estado del desarrollo de un juvenil antes de adquirir los caracteres del adulto, concepto que no hay ningún inconveniente en aplicar a los cefalópodos. En sentido estricto, larva define un estadio juvenil que se transforma en adulto a través de una metamorfosis, la cual se caracteriza por dos procesos principales: obliteración de las partes larvarias del animal y formación de ciertas partes del adulto a partir de rudimentos que permanecían en esbozo. Analizando con detenimiento las transformaciones que experimentan los juveniles hasta llegar a adultos se comprueba que estas condiciones no se cumplen, por lo que no debería hablarse en puridad de estados larvarios de los cefalópodos. Hecha esta precisión, emplearemos aquí el término larva como una abreviatura práctica, igual que se hace en otros grupos taxonómicos, por ejemplo en los peces.

Las larvas de los ommastréfidos, conocidas por *Rhynchoteuthion*, carecen del par de brazos ventrales, e incluso del tercer par en algunas especies; poseen rostro o trompa con ventosas en la parte apical, que durante el crecimiento se escinde longitudinalmente dando lugar a los tentáculos. Aunque se ignora el comportamiento trófico y la dieta de las larvas *Rhynchoteuthion*, su típica configuración sugiere un régimen y comportamiento alimentarios diferentes de los que muestran los juveniles, dotados de tentáculos y brazos perfectamente desarrollados, y también de los adultos, que son activos depredadores. Verosímilmente, una probóscide con ventosas en el ápice puede ser más eficaz que los tentáculos a la hora de capturar presas lentas. Si la eclosión de esta especie estuviese sincronizada con la de otras que fuesen lentas, la gran abundancia de presas de este tipo podría haber sido un factor decisivo en la selección del carácter.

Esta hipótesis adquiere mayor probabilidad cuando se tiene en cuenta este otro dato: las larvas de *Illex illecebrosus* se capturan en el Atlántico noroccidental en masas de agua de origen tropical, cuyos organismos indicadores son moluscos pterópodos de las especies *Limacina trochiformis, Cresis virgula* y *Cavolinia longirostris*, totalmente pelágicos y cuyas larvas se desenvuelven con movimientos lentos. La brevedad del proceso que conduce a la aparición de tentáculos, y con ello a un cambio de la dieta, indica además que las larvas *Rhynchoteuthion* se aprovechan de grandes, aunque breves, aumentos de biomasa de larvas de pterópodos, que ocurren en los núcleos de los típicos remolinos que forma la corriente de agua templada del Golfo, en su convergencia con las masas de agua de origen boreal. Se trata de zonas con una producción primaria destacadamente alta.

Otro ejemplo de un fenómeno similar es el de la pota del Atlántico norte, *Todarodes sagittatus*, que tiene su área

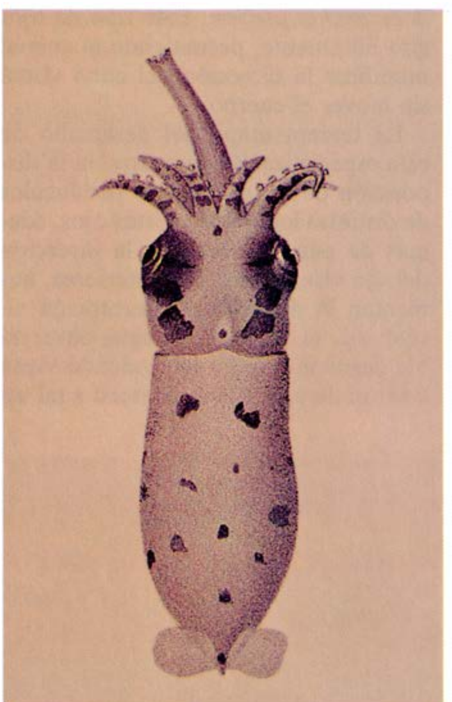
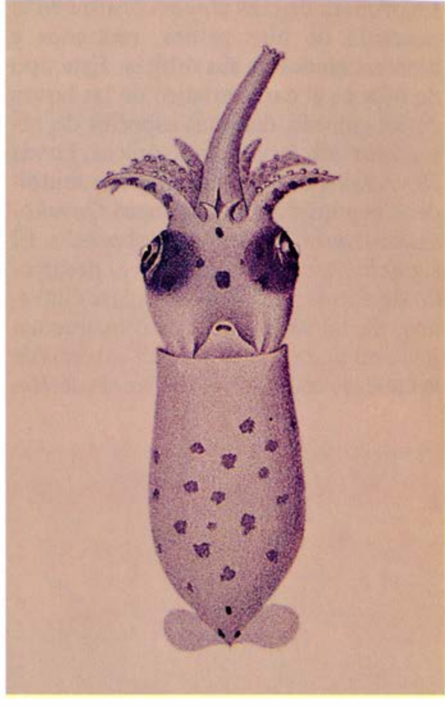


3. CUADRO DIAGRAMATICO CON LA CLASIFICACION de la clase Cephalopoda. Unicamente se han incorporado los grupos taxonómicos que se citan en el texto. El orden Sepioidea consta de cinco familias. El suborden Myopsida está formado por dos familias; Oegopsida, por 23; Cirrata, por tres e Incirrata por nueve familias. Así, los cefalópodos están subdivididos en 44 familias que comprenden cerca de 800 especies. No obstante, la clasificación de esta clase dista todavía de ser definitiva, estando varios grupos en revisión. Los dibujos representan el tipo más característico de cada familia citada.

de puesta en la zona comprendida entre el noroeste de Irlanda y las islas Azores y Canarias. Esta especie invade algunos años las costas de Noruega y zonas adyacentes, invasiones que coinciden con una gran abundancia de salpas (*Salpa fusiformis*), organismos indicadores, a su vez, de agua de origen atlántico. La zona de procedencia de las salpas coincide con el área de puesta de *Todarodes sagittatus*, donde la productividad es bastante elevada especialmente en primavera. La principal época de puesta de *Todarodes sagittatus* abarca desde diciembre hasta febrero; la eclosión de las larvas viene a coincidir con la época de mayor abundancia de biomasa zooplanctónica.

Otra característica adaptativa de las larvas de los cefalópodos que interesa resaltar tiene que ver con la forma de las mandíbulas. En las larvas planctónicas de calamar de las especies *Loligo vulgaris* y *L. pealei* se ha observado que el borde anterior de la mandíbula inferior es dentado; presenta dos dientes particularmente bien definidos, uno a cada lado de una incisión central. La mandíbula superior carece de dientes. Parecidas a las mandíbulas de *Loligo* son las de las larvas *Rhynchoteuthion*. Aunque no se ha realizado todavía una investigación exhaustiva, todo indica que esta configuración de la mandíbula aparece en los teutoideos miópsidos y en los oegópsidos. En numerosas especies de Octópodos incirrados, como *Octopus vulgaris, O. briareus, O. joubini, O. maya, Eledone cirrhosa, Argonauta argo* y *Tremoctopus violaceus*, se ha comprobado que el borde anterior de ambas mandíbulas es dentado. Los dientes, bastante regulares, muestran una forma cónica. Esta dentición es más importante en las especies que presentan larvas planctónicas *–O. vulgaris, Eledone cirrhosa, Argonauta argo* y *Tremoctopus violaceus–* que en aquellas cuya descendencia es bentónica, como *O. briareus* y *O. maya*, en que los bordes, más reforzados, se parecen también más a los bordes cortantes y lisos de la mandíbula de los adultos. *Octopus joubini*, de larvas bentónicas, presenta un estado intermedio entre estos dos.

En todos los individuos recién nacidos pertenecientes al orden Sepioidea, tanto si se trata de especies que nacen con tamaño relativamente grande como *Sepia officinalis*, o larvas de 1,6 milímetros de longitud del manto pertenecientes a la especie *Heteroteuthis dispar* y recogidas con el plancton, las mandíbulas son similares a las de los adultos; carecen de dentición.

4. LARVA *RHYNCHOTEUTHIS* en visión dorsal y ventral. Estas larvas son típicas de la familia Ommastrephidae. La trompa que se ve dará lugar a los tentáculos al escindirse longitudinalmente durante el crecimiento. En el ápice de la probóscide se forman seis u ocho ventosas de diverso tamaño. Se distinguen los cortos brazos ventrales y el comienzo de la escisión de la trompa que avanza desde la zona basal.

La estructura de las mandíbulas guarda una estrecha relación con el tipo de vida de las larvas. Todas las larvas de teutoideos estudiadas tienen mandíbulas dentadas adaptadas a su vida planctónica. Los sepioideos, cuyas larvas son bentónicas, a excepción de *Heteroteuthis dispar* y *Spirula spirula*, mantienen, desde el nacimiento, mandíbulas similares a las de los adultos. En estos dos géneros de larvas planctónicas, la existencia de mandíbulas sin dientes quizá se deba a un cambio reciente de hábitat.

En los octópodos se observa una gradación de la dentición de las mandíbulas de las larvas, desde las dentadas en especies con larvas planctónicas, así *O. vulgaris*, hasta las lisas de especies con larvas bentónicas, *O. maya*, por ejemplo.

Aunque todavía no se han realizado estudios detallados sobre el desarrollo postembrionario de las mandíbulas de los cefalópodos, los primeros datos obtenidos por Boletzky en *Loligo vulgaris* muestran que la dentición desaparece 45 días después de la eclosión.

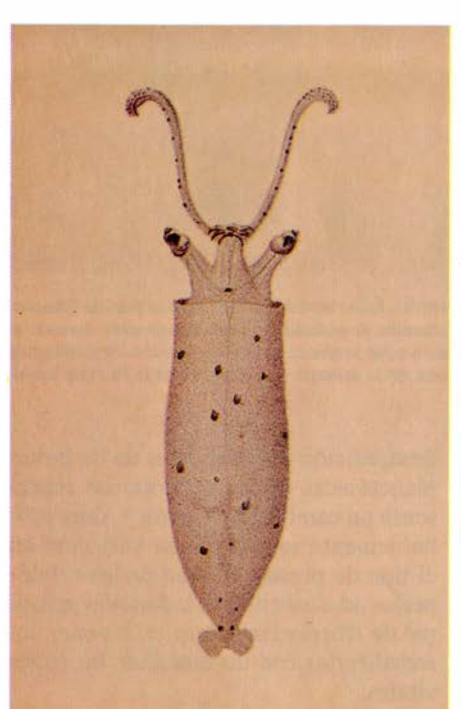
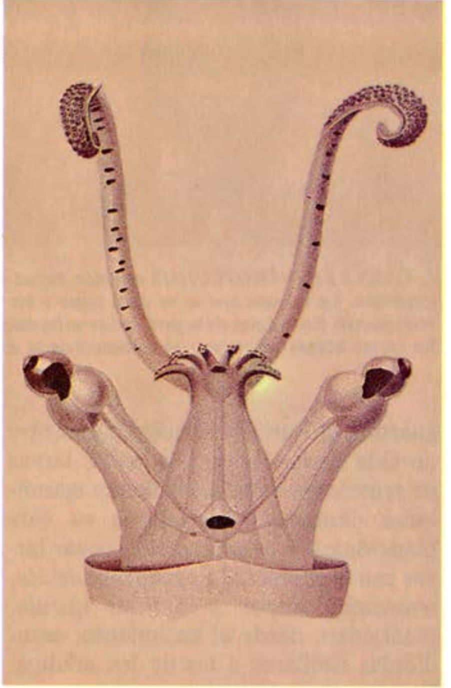
La presencia de mandíbulas dentadas es frecuente en organismos planctónicos de diferentes grupos. Puede considerarse una adaptación al tipo de régimen alimentario macrófago, en que las presas son capturadas, sujetadas por apéndices y posteriormente desgarradas por los dientes de las mandíbulas antes de ser deglutidas. El progresivo reforzamiento de las mandíbulas y la desaparición de los dientes de las larvas planctónicas de los cefalópodos representa un cambio de función y, muy probablemente también, una variación en el tipo de presas. El pico de los cefalópodos adultos cumple la función principal de trocear las presas, y, a veces, inmovilizarlas con un mordisco en zonas vitales.

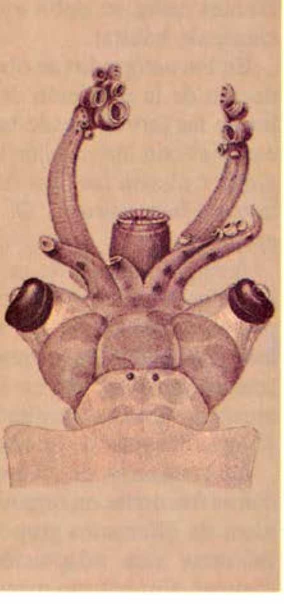
El tamaño y forma de los ojos de las larvas planctónicas de los cefalópodos, especialmente de los teutoideos oegópsidos, se distinguen por su amplia variación. Considerando que las causas de mortalidad más importantes para las larvas son el hambre, la depredación y las propias interacciones, y que tanto la alimentación como la supervivencia a la depredación dependen fundamentalmente de las adaptaciones visuales, es de esperar que aquellas especializaciones oculares que favorezcan una mayor eficacia visual se hayan potenciado en el transcurso de la evolución. En larvas de peces se han encontrado adaptaciones parecidas; se ha demostrado que las dos principales limitaciones para la alimentación de las mismas son la distancia de percepción y la abertura del campo visual. Weihs y Moser han sugerido que el aumento del campo visual en las larvas de los peces ha seguido un curso progresivo, a través de varias etapas.

Un modelo parecido al de los peces puede encontrarse en los cefalópodos.

La primera de esas etapas consiste en la posesión de ojos sésiles, redondos y bien encajados en sus órbitas. Este tipo de ojos es el característico de las larvas de loligínidos, distintas especies de octópodos con larvas planctónicas, larvas *Rhynchoteution* y otras larvas de teutoideos oegópsidos, por ejemplo *Onychoteuthis banksii* y *Gonatopsis borealis*. El siguiente estadio consiste en el desarrollo de cortos pedúnculos oculares internos, de tal forma que los ojos queden justo en línea con el margen externo de la cabeza, como los de las larvas de *Helicocranchia pfefferi*. Este tipo de ojos gira libremente, permitiendo al animal modificar la dirección del cono visual sin mover el cuerpo.

La tercera etapa del desarrollo de esta especialización se centra en la disposición de los ojos sobre pedúnculos de distintas longitudes. Estos ojos, además de permitir cambiar la dirección del eje visual como los anteriores, aumentan la distancia de percepción visual; así, el volumen de agua observable desde un punto determinado viene a ser el de una esfera. Merced a tal especialización, se detecta la presa desde mayor distancia y con un mínimo de movimiento del animal, lo que supone menor gasto energético. Al no tener que moverse, el animal no produce perturbaciones en el agua que podrían atraer a los depredadores. Los ojos pedunculados son frecuentes en las larvas de los cránquidos –piénsese en *Theuthowenia megalops* y *Pyrgopsis pacifica*, especies de pedúnculos cortos, en *Bathothauma lyromma* y *Sandalops melancholicus*, de pedúnculos mucho más largos, de un tamaño aproximado a la mitad de la longitud del manto del animal–. El hecho de que este tipo de ojos, sobre todo los de pedúnculo corto, haya aparecido también en otras especies de diferentes familias, como por ejemplo en *Brachioteuthis riisei*, en distintas larvas de chiroteútidos, octopoteutidos y otros, nos habla de su eficacia adaptativa.

5. TIPOS DE OJOS PEDUNCULADOS de las larvas. Las dos fotografías superiores y la izquierda de la serie inferior pertenecen a *Corynomma speculator* (familia Cranchiidae); la central, a *Bathotauma lyromma*, de la misma familia. Por último, en la derecha se muestra el detalle de la cabeza y de los ojos de *Octopoteuthis sicula* (familia Octopodoteuthidae). Estas larvas tienen ojos dispuestos sobre pedúnculos de distintas longitudes, que, además de permitir cambiar la dirección del eje visual, aumentan la distancia de percepción. Merced a especialización detectan la presa a mayor distancia y sin apenas movimiento.

En *Bathotauma lyromma* los pedúnculos llegan a medir hasta 27 milímetros de longitud. Weihs y Moser establecieron una fórmula que relaciona la percepción de las dos clases de ojos: $Vs/Va = (1 + a/d)^3$, donde *Vs* es el volumen de percepción de un ojo pedunculado; *Va*, el de un ojo del tipo segundo; *a*, la longitud del pedúnculo y, *d*, la distancia de percepción visual. Dicha percepción se considera de 10 milímetros en los peces. Según la fórmula, el volumen de percepción de los ojos pedunculados de *B. lyromma* es 50 veces mayor que en *H. pfefferi.* Teniendo en cuenta que en ojos que no pueden modificar la dirección del cono visual sólo es observable un 30 por ciento de su volumen de percepción, los ojos de *B. lyromma* con respecto a los del primer tipo tienen una ganancia en volumen de percepción visual de unas 600 veces.

Con este aumento de volumen observable se mejora el balance entre energía adquirida por la alimentación y energía gastada en búsqueda, relación que importa sobremanera en áreas de baja concentración de alimento. Los pedúnculos desaparecen con el crecimiento, debido, quizás, a la diferente estrategia que siguen las larvas y los adultos en relación con sus presas y depredadores. En el primer caso, la caza se realiza al acecho y se evita a los depredadores pasando desapercibido, mientras que los adultos persiguen a sus presas y escapan de sus depredadores. En las larvas se prima, pues, la percepción visual y la inmovilidad; en los adultos, la velocidad.

Las larvas de los calamares oceánicos pertenecientes al género *Chiroteuthis*

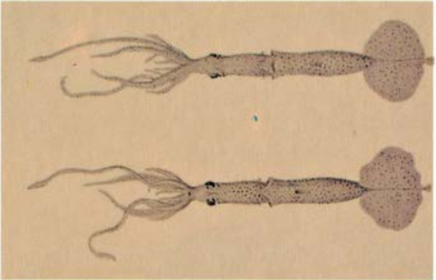
poseen una morfología muy curiosa y sensiblemente diferente de la que poseen los adultos. Los individuos de 0,5 a 6,5 centímetros de longitud del manto, capturados en el plancton, presentan brazos poco desarrollados, un cuello larguísimo, de longitud similar a la del manto, el cual acaba en una cola puntiaguda bastante larga y al que se adosan dos aletas laterales redondeadas. También las larvas de *Brachioteuthis riisei*, especie perteneciente a la familia Brachioteuthidae, se distinguen por su cuello alargado. La aparición de esta peculiar morfología larvaria, de manera independiente en familias distintas, es una prueba de que se trata de una especialización adaptativa.

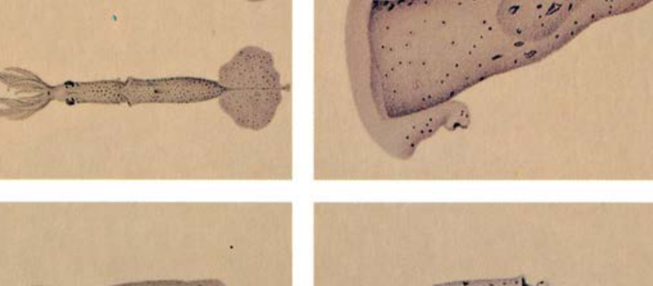
Aunque actualmente se desconoce casi todo lo referente al comportamiento de este tipo de larvas, es posible avanzar algunas hipótesis para explicar sus peculiaridades morfológicas. En primer lugar, el mantenimiento de la cabeza sobre su cuello largo con gran libertad de movimiento con respecto al cuerpo puede contemplarse como posesión de una atalaya donde están situados los ojos. Los movimientos del cuello significarían un sensible aumento del campo visual y del volumen de agua inspeccionado sin moverse la larva de sitio, lo cual, como ya hemos visto, constituye una apreciable ventaja adaptativa frente a presas y depredadores. El cuello largo vendría a desempeñar una función análoga a los pedúnculos oculares, habiéndose resuelto un mismo problema por dos procedimientos distintos.

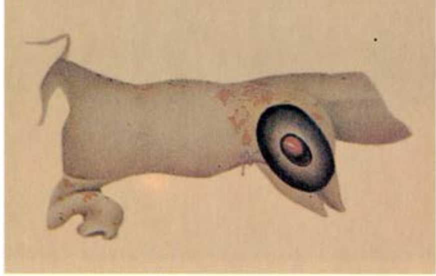
El cuello y la cola pueden servirles también para los desplazamientos verticales dentro de la masa de agua. El primero actuaría entonces de flotador y la cola de estabilizador, especialmente en los movimientos descendentes. La variación de flotabilidad se debería a cambios de volumen del cuello, bien actuando como un fuelle o introduciéndose dentro del manto.

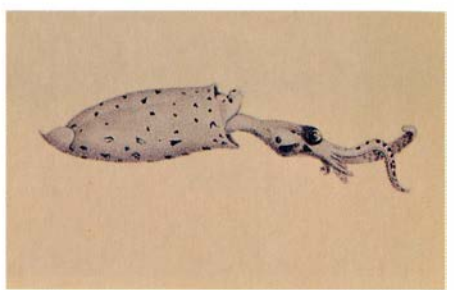
Otro tipo de especialización cuya función algunos autores han relacionado con el retardo de la velocidad de caída de las larvas planctónicas, o como medio de defensa, son los órganos de Kölliker. Aparecen en todos los jóvenes octópodos incirrados conocidos y semejan pequeñas brochas distribuidas por toda la superficie del cuerpo.

Los cefalópodos basan su locomoción en la natación a reacción provocada por la actividad del complejo manto-sifón, y también de las aletas; aunque hay grupos, los octópodos sin cirros entre ellos, que carecen de aletas. Desde su nacimiento, la mayoría de las larvas tienen perfectamente de-









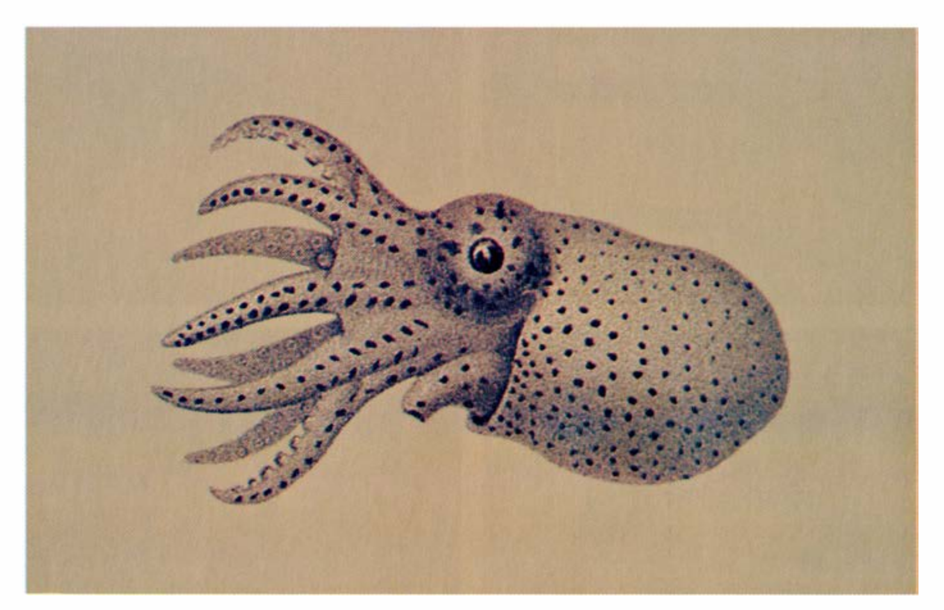
6. LARVAS CON EL TIPICO CUELLO LARGO, cuya longitud llega a ser igual a la del manto. Salvo la inferior derecha, que pertenece a la especie *Brachioteuthis riisei*, todas son *Doratopsis*. Esta peculiaridad morfológica puede tener varias finalidades, una de ellas sería análoga a los pedúnculos oculares y otra podría estar relacionada con los desplazamientos verticales, actuando el cuello como flotador.

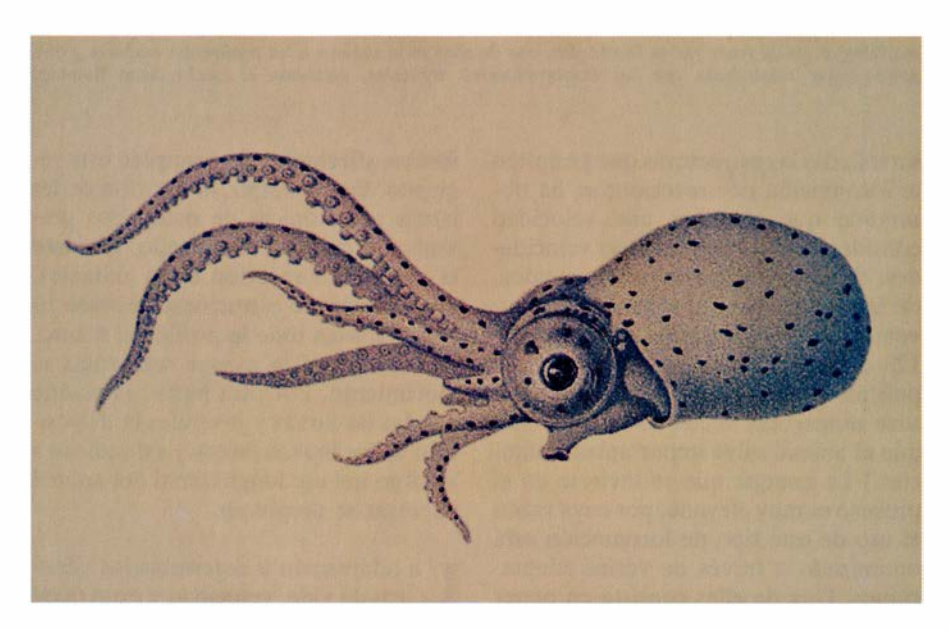
sarrolladas las estructuras que permiten la locomoción por reacción; se ha observado que alcanzan una velocidad considerable. Se han medido velocidades, de *Loligo vulgaris* recién nacidos, de 16 centímetros por segundo; y en juveniles de tres gramos de peso hasta de 128 centímetros por segundo. (La expulsión del chorro de agua puede repetirse numerosas veces, lo que permite que el animal salve importantes distancias.) La energía que se invierte en el proceso es muy elevada, por cuya razón el uso de este tipo de locomoción está optimizado a través de varias adaptaciones. Una de ellas consiste en poner oblicuo el cuerpo, durante la natación, cabeza abajo y formando un ángulo de unos 45 grados. Una disposición similar se ha observado en los animales que salen proyectados del agua, como un pez volador o un salmón, que alcanzan mayores distancias para un mismo impulso muscular si salen de la interfase agua-aire con un ángulo de 45 grados. En estos movimientos rápidos, las aletas desempeñan un importante papel como estabilizadores; los individuos que han sufrido algún daño en las mismas no consiguen mantener la posición correcta durante un desplazamiento a reacción. Las larvas planctónicas de los octópodos, que carecen de aletas, se comportan igual que las larvas de los teutoideos, pero su mecanismo estabilizador se desconoce todavía, aunque quizás utilicen para ello los brazos.

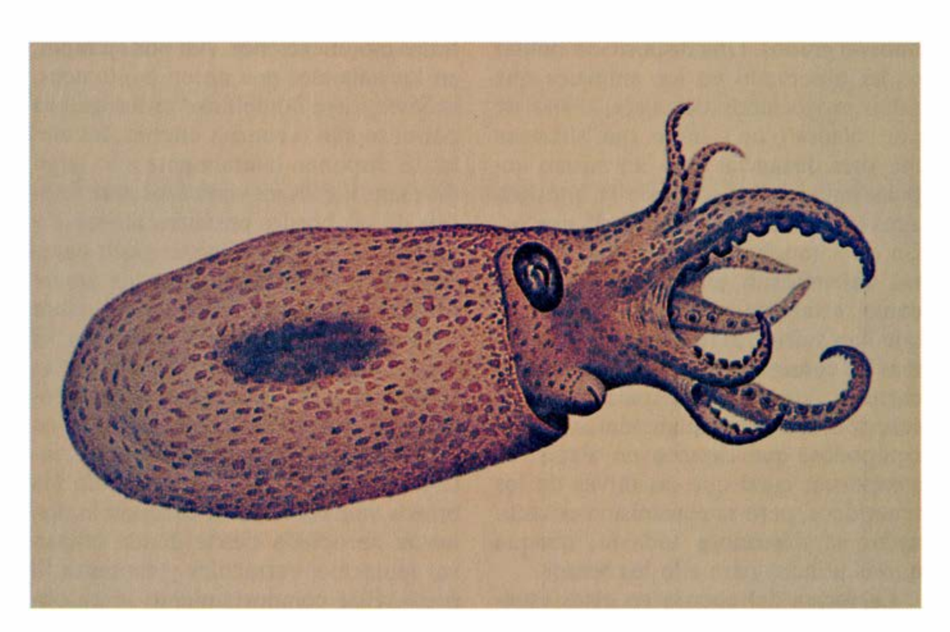
La forma del cuerpo en estas especies veloces debe aproximarse a la sección que ofrezca menor resistencia al desplazamiento. Muchas larvas tienen formas aflechadas que cumplen este requisito. Sin embargo, la mayoría de las larvas planctónicas de octópodos presentan el cuerpo redondeado. Durante la natación a reacción estos animales, que carecen de estructuras internas rígidas, ahusan todo lo posible el manto, ofreciendo así la menor resistencia al movimiento. Por otra parte, es común a todas las larvas y juveniles la disposición de los brazos juntos y extendidos a lo largo del eje longitudinal del animal mientras se desplazan.

La adaptación a determinados géneros de vida, velocidad y exigencias de maniobrabilidad en espacios reducidos ha impuesto en los cefalópodos distintas modificaciones. Así por ejemplo, en los animales que nacen bentónicos, la *Sepia*, cuya flotabilidad está regulada por el sepión o concha interna, las aletas se disponen lateralmente a lo largo del manto, y los movimientos ondulatorios de sus bordes permiten moverse y maniobrar con un mínimo gasto energético y gran facilidad. Pueden atacar sin usar la locomoción a reacción, sobre todo cuando las presas están en el fondo; en ese caso, las jóvenes sepias se acercan a su presa con el cuerpo horizontal, los brazos extendidos y en contacto con el fondo, avanzando lentas, casi caminando con ayuda de los brazos ventrales, hasta alcanzar la distancia apropiada desde donde lanzan sus tentáculos retráctiles y capturan la presa. Este comportamiento se ha observado también en sepiólidos juveniles, particularmente en *Rossia*.

El problema de ofrecer la mayor re-







7. LARVAS DE OCTOPODOS INCIRRADOS. De arriba abajo: *Octopus* sp, *Tremoctopus* sp, *Bolitaena diaphana*. Son de vida planctónica. Las diferencias con los adultos, pequeñas, se limitan casi al tamaño corporal. La falta de caracteres adaptativos especiales las señalan como larvas de tipo primitivo.

sistencia posible al arrastre de la corriente lo han resuelto las larvas de maneras distintas. En los recién nacidos de *Sepia officinalis*, la musculatura tegumentaria de la cara ventral del manto y de la región inferior de los brazos actúa como una ventosa, para adherirse a los fondos, resistiendo los embates del agua. Esta especialización transitoria es exclusiva de los cefalópodos que tienen el cuerpo rígido debido a la concha interna o sepión. Los sepiólidos y los octopódidos, que carecen de una concha interna desarrollada, aseguran su fijación al fondo mediante las ventosas de los brazos; la superficie de fijación es relativamente mayor que en las sepias y su cuerpo flexible ofrece menos resistencia a la corriente.

En las larvas de cefalópodos existen otros mecanismos que favorecen su supervivencia. Uno de ellos es la posibilidad de lanzar nubes de tinta. Prácticamente todas las larvas de los cefalópodos decápodos, con excepción de *Neorossia caroli*, y la mayoría de los octópodos tienen bolsa de la tinta funcional desde su nacimiento. En los octópodos batipelágicos adultos se observa una reducción de dicha bolsa, relacionada seguramente con la escasa intensidad luminosa de las grandes profundidades en que viven. Se desconoce si sus larvas presentan también este órgano en regresión.

Las pequeñas larvas planctónicas suelen portar escasos cromatóforos, de tamaño relativamente grande; de ahí la sencillez del repertorio de cambios de coloración, aparición de manchas epidérmicas y extensión de las mismas que las caracteriza. Sin embargo, los recién nacidos que llevan desde el comienzo vida bentónica gozan ya de capacidad para exhibir un repertorio cromático tan variado como el de los adultos, que les capacita para la mímesis en los distintos fondos en que viven. Este tipo de cripsis u ocultación no se limita a cambios de coloración, sino que se amplía con la adquisición de aspectos, posturas y formas diferentes, que se reducen casi por entero a las modificaciones de la textura de la piel; éstas confieren al animal apariencias extrañas, idóneas para un camuflaje eficaz.

La presencia de mecanismos complejos de defensa basados en el mimetismo en los cefalópodos bentónicos recién nacidos, la mayor sencillez de estos mecanismos en las larvas planctónicas y en los adultos pelágicos y la riqueza del repertorio de respuestas cromáticas en las larvas planctónicas de algunos octópodos, a medida que crecen y comienzan a vivir en el fondo, dan fe

de que en los ecosistemas más estables, considerando por tales a los bentónicos frente a los pelágicos y planctónicos, hay una reducción efectiva del flujo de energía relativo y de la tasa de renovación en general, como ha señalado Ramón Margalef; la presencia de relaciones defensivas muy elaboradas representa la posibilidad de una mayor tasa de supervivencia. En los cefalópodos con juveniles bentónicos, se observa una reducción de la producción de huevos y un aumento de tamaño de los mismos; así una hembra de *Sepia officinalis* pone de 200 a 500 huevos cuyo tamaño varía entre 9 y 4,5 milímetros, lo que indica una alta supervivencia de los juveniles en las primeras fases de su vida. Por el contrario, en especies bentónicas, como *Octopus vulgaris* cuyas larvas desarrollan una vida planctónica, la producción de huevos es más elevada, oscilando alrededor de 200.000 huevos por puesta. La tasa de reproducción de *O. vulgaris* corrobora que la tasa de mortalidad de las larvas de esta especie es más elevada que en *Sepia*.

Prueba evidente de que las larvas de algunas especies de cefalópodos difieren bastante de las formas adultas nos la ofrece la historia de su sistemática: en varias ocasiones, dichas larvas se han descrito como auténticas especies. Las larvas de los ommastréfidos fueron clasificadas por Carl Chun, en 1903, como un nuevo género, *Rhynchoteuthis*, que se cambió posteriormente por el de *Rhynchoteuthion* porque el anterior había servido para catalogar un fósil. En 1910, el propio Chun comprobó que esos organismos eran larvas de ommastréfidos; el género quedó invalidado, aunque todavía se sigue utilizando para designar estas larvas. Algo similar ocurrió con los primeros estados larvarios de los chiroteutidos, que recibieron diferentes nombres como *Planctoteuthis, Cheiroteuthis* y *Doratopsis*, según su estado de desarrollo. En 1899 Ficalbi sugirió que la especie descrita como *Doratopsis vermicularis* no era sino un estadio juvenil de *Chiroteuthis verany*, lo que actualmente está plenamente comprobado y unánimemente admitido, tras treinta años largos de observaciones y controversia. El problema del *Macrotritopus* es también muy ilustrativo. Se dio este nombre genérico a un pulpito pelágico cuyo tercer par de brazos es mucho mayor que los demás. Durante más de sesenta años se le consideró un género o un subgénero de *Octopus*. En 1954 se comenzó a sospechar que podría ser una fase juvenil de cualquiera de estas cuatro especies: *Octopus vulgaris, O. macropus, Scaeurgus unicirrhus* y *Pteroctopus tetracirrhus*. En ese mismo año, cuando se conocieron las larvas de las dos primeras especies, se pensó que también podría serlo de *O. defilippi*, descartándose la posibilidad de que fuese una larva de *P. tetracirrhus*. La prueba directa de que *Macrotritopus* corresponde a una fase del desarrollo de *O. defilippi* se obtuvo en 1979, cuando Hulet y su equipo consiguieron completar el desarrollo hasta el estado adulto de dos ejemplares *Macrotritopus*, capturados por la noche a 50 metros de profundidad.

El conocimiento de las larvas resulta fundamental para explotar de forma juiciosa las pesquerías de cefalópodos. En la pesquería de *Todarodes pacificus* del mar del Japón, de la que se extraen anualmente más de medio millón de toneladas, estos estudios se vienen realizando desde hace bastantes años. Más modernos son los esfuerzos por conocer estos aspectos en la pesquería de *Illex illecebrosus* de las costas de Canadá. La investigación arranca del conocimiento taxonómico de las larvas. En el Atlántico occidental se describieron en 1978 tres tipos distintos de larvas *Rhynchoteution*. Después de varios trabajos, entre los que tuvieron especial importancia los experimentos realizados en el laboratorio, que permitieron obtener descendencia de puestas de esta especie, se identificaron las etapas del desarrollo de las larvas de *Illex illecebrosus*, distinguiéndose de *Illex oxigonius, Ornithoteuthis antillarum* y de varias especies del género *Ommastrephes*. También son de gran importancia las campañas científicas encaminadas a descubrir la abundancia y distribución de las larvas que permitan conocer la época y la localización de la freza, vías de migración, existencia de diferentes unidades de población, abundancia anual y ritmo de crecimiento en las primeras fases.

El estudio de las larvas de los cefalópodos está en sus comienzos. Como el interés por estos organismos va en aumento, el número de campañas científicas se ha incrementado, y se ha avanzado considerablemente en los estudios en acuarios experimentales, es de esperar que en los próximos años muchos de los aspectos morfológicos y funcionales que todavía se desconocen, algunos de los cuales se han tratado en este trabajo, se revelarán con toda claridad, contribuyéndose así al conocimiento de estos organismos perfectamente adaptados al medio marino, cuyos adultos constituyen una importante fuente de alimento que debe explotarse de una manera racional.

Io

El mundo volcánico más activo del sistema solar es un satélite de Júpiter cuyas erupciones en géyser y enormes ríos de lava delatan la existencia de exóticos fluidos de origen volcánico

Torrence V. Johnson y Laurence A. Soderblom

En la moderna era de exploración planetaria, Io, satélite de Júpiter descubierto por Galileo hace 400 años, ha resultado ser uno de los mundos más extraños del sistema solar. En primer lugar, hacia el decenio de 1960, los astrónomos descubrieron que, de alguna manera, Io modula los intensos brotes de radioondas emitidos por Júpiter. Además, instrumentos sensibles apuntados hacia Io desde la Tierra confirmaron que el objeto es tan reflector como la nieve fresca, a pesar de su color amarillo anaranjado. Esto abrió ciertos interrogantes sobre la naturaleza de la superficie de Io. Desde el decenio de 1970 se sabe que inyecta grandes nubes de iones y de átomos neutros en la magnetosfera de Júpiter, o región donde las partículas eléctricamente cargadas quedan aprisionadas por el campo magnético del planeta. Finalmente, los vehículos espaciales penetraron en el sistema joviano de satélites. Las pasadas de los *Voyager 1* y *Voyager 2* en 1979 proporcionaron la primera vista cercana de Io.

El Io que conocemos hoy nos parece todavía más extraño. Calentado por fuerzas de marea ejercidas por Júpiter, constituye, por lo que se sabe, el cuerpo volcánicamente más activo del sistema solar. Su actividad supera sin duda alguna la terrestre. Grandes géyseres volcánicos envían allí penachos eruptivos hasta alturas de cientos de kilómetros, y la materia que escapa de la tenue atmósfera volcánica de Io, a través de procesos todavía no bien conocidos, domina la magnetosfera joviana, controlando así el ritmo al que Júpiter radia energía e influyendo probablemente en fenómenos tan distintos como las auroras de Júpiter y los brotes de radioemisiones dimanantes del planeta y sus satélites. Los descubrimientos hechos por los vehículos espaciales Voyager han provocado un aluvión de preguntas. No cabe pensar aún en ningún artículo definitivo, y este nuestro se limita a ser un informe sobre la situación actual de la investigación acerca de tan notable objeto celeste.

Io se cuenta entre los satélites "normales" de Júpiter; es decir, se mueve en órbitas esencialmente circulares en el mismo plano que el ecuador; formado, por tanto, alrededor de Júpiter de manera muy parecida a como los planetas se constituyeron alrededor del Sol. A una distancia orbital de 350.000 kilómetros (aproximadamente la misma que la distancia de la Luna a la Tierra), es el más interno de los cuatro grandes satélites jovianos que descubrió Galileo. Io y Europa, el siguiente hacia fuera, vienen a ser del tamaño de la Luna terrestre; tienen una densidad parecida a la que poseen las rocas: 3,5 gramos por centímetro cúbico y 3 gramos por centímetro cúbico, respectivamente. Ganimedes y Calixto, los más externos de los cuatro, son mayores (del tamaño aproximado de Mercurio) y tienen una densidad apreciablemente más baja, del orden de dos gramos por centímetro cúbico. Al parecer están formados a partes iguales por hielo y roca, silicatos principalmente. La variación de densidad entre los cuatro cuerpos se ha calificado con frecuencia de curso suave, aunque se acerca más a una distribución bimodal. La masa deducida para los silicatos en Ganimedes y Calixto es aproximadamente igual a la masa total de Io y Europa; así, los cuatro pueden considerarse similares con la salvedad de que en los dos más internos ha desaparecido su cupo de materias volátiles, agua sobre todo. (Se sabe que Europa está cubierta de hielo, pero su densidad global indica que esta cobertura es sólo un barniz de espesor inferior a una vigésima parte del radio del satélite.)

La variación de densidad suele atribuirse al efecto producido por Júpiter sobre las primitivas condiciones en su vecindad. Aun hoy, el planeta radia en longitudes de onda del infrarrojo aproximadamente el doble de la energía que recibe del Sol; se ha estado, pues, enfriando a partir de un primitivo estado luminoso. De aquí que los modelos de su primitiva historia (basados en una física semejante a la de las estrellas de baja masa) nos hablen de un planeta constituido en importante fuente local de energía durante varios cientos de millones de años. Los modelos desarrollados por James B. Pollack, del Centro de Investigación Ames de la Administración Nacional de Aeronáutica y del Espacio (NASA), y sus colaboradores sugieren que, a lo largo de un período de unos 100 millones de años, la temperatura y la presión en la posición orbital actual de Io estaban dominadas por Júpiter, no por el Sol. En efecto, un cuerpo situado en la órbita de Io en el primer millón de años hubiera recibido de Júpiter más energía que la cantidad que la Tierra recibe del Sol ahora.

Es probable, por tanto, que la composición de los satélites normales de Jú-

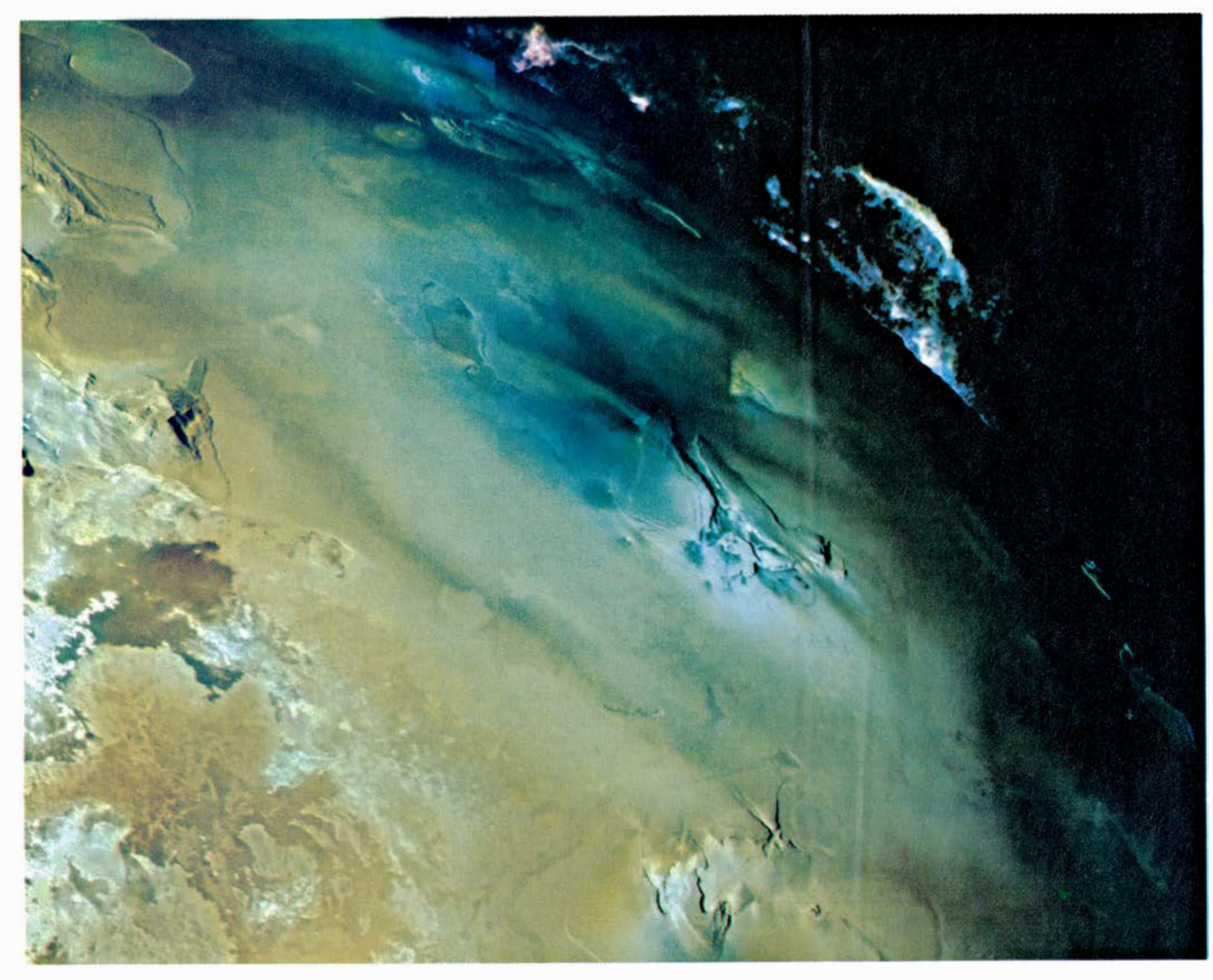
1. HEMISFERIO SUR DE IO, fotografiado en 1979 por una cámara con gran angular montada en el vehículo espacial *Voyager 1*. El campo de visión abarca aproximadamente 2,5 millones de kilómetros cuadrados, es decir, un 6 por ciento de la superficie de Io; un centímetro representa unos 50 kilómetros. Se ven varias clases de accidentes. Por su número sobresalen las calderas (cráteres volcánicos colapsados). Algunas están rodeadas de marcas difusas, que probablemente representan el pintado de la superficie por géyseres sulfurosos. Otras están rodeadas por probables ríos eruptivos. Hay todavía otras que parecen llenas de material eruptivo; la caldera romboidal con manchas rojas y negras es un ejemplo. Las manchas blancas que motean la superficie se creen debidas a dióxido de azufre gaseoso y escarchado que escapa a través de escarpaduras y fracturas. Se supone que la montaña que se distingue en la parte inferior izquierda constituye un afloramiento de la corteza de silicatos de Io; se le atribuyen unos 10 kilómetros de altura. La imagen se ha construido con un ordenador del Servicio de Inspeccción Geológica de los Estados Unidos de tal manera que se nos ofrezcan los accidentes como si estuviéramos sobrevolando el paisaje.



piter estuviese condicionada por su relativa proximidad al planeta. Según este punto de vista, Io y Europa son cuerpos rocosos porque las condiciones cerca de Júpiter eran demasiado calurosas para permitir a los satélites en formación retener cantidades importantes de agua. Por contra, Ganimedes y Calixto, que sí retuvieron agua, se convirtieron en mezclas de hielo y roca y adquirieron un tamaño mayor que Io y Europa. Adviértase que las densidades de Ganimedes y Calixto son, aproximadamente, las que cabría esperar en la hipótesis de que un gas de la composición del Sol se enfriase hasta la temperatura de condensación del hielo.

Esta sencilla interpretación no puede explicar todos los datos. Por un lado, los cuerpos del sistema solar recibieron un ataque intenso de planetésimos en un período llamado de bombardeo pesado, que acabó hace unos cuatro mil millones de años. Los planetésimos que caían en el campo gravitatorio de Júpiter debieron ser grandemente acelerados; lo que, para Io y Europa, supuso un vapuleo más completo que el que sufrieron los satélites jovianos más alejados. Además, la intensa actividad volcánica de Io ha alterado su propio contenido en sustancias volátiles. Pese a ello, diríase que los satélites normales de Júpiter, y más en concreto Io, se formaron en procesos semejantes a los que produjeron los planetas. Júpiter y sus satélites constituyen así un sistema solar en miniatura en un sentido más profundo que el puramente geométrico.

La superficie de Io que nos muestran las imágenes del Voyager está dominada por la actividad volcánica, mucho más allá de lo esperado. Por la siguiente razón: se cree que las influencias normales sobre la evolución térmica (y por tanto sobre el vulcanismo) en un cuerpo planetario son, en primer lugar, el calentamiento del cuerpo por liberación de la energía potencial gravitatoria en su aglutinación y también por desintegración de isótopos radiactivos de corta vida, tales como el aluminio 26; en segundo lugar, la gradual acumulación de calor debida a la desintegración de isótopos radiactivos de vida larga, principalmente uranio, torio y potasio; y, finalmente, un enfriamiento gradual a medida que se van agotando las fuentes radiactivas y se pierde calor a través de la superficie del cuerpo en virtud de la conducción y la convección. Semejante sucesión favorece la actividad volcánica en cuerpos sólidos grandes, los cuales tienen una relación de volu-



2. PELE, el primer volcán descubierto en Io, es también la mayor erupción en forma de géyser observada. El penacho del géyser sobresale en el confín del satélite; asciende a una altura de unos 300 kilómetros, y ha depositado sobre la superficie un grupo de anillos concéntricos amarillos y castaños. El anillo más externo tiene en promedio 1400 kilómetros de diámetro. En el centro del depósito yace la fuente del géyser: un complejo de colinas con un valle central. Las marcas y los ríos son prueba de anteriores erupciones. Esta vista de Pele fue tomada por el *Voyager 1*; se trata de un mosaico producido mediante una técnica desarrollada por Alfred S. McEwen, de la Inspección Geológica, en la que las imágenes de alta resolución aportan el detalle espacial y las imágenes de baja resolución suministran datos de color. Cuando el satélite *Voyager 2* llegó cuatro meses después, había cesado la actividad volcánica en Pele.

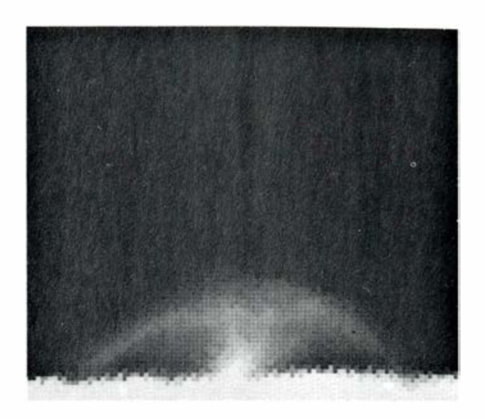
men a área de su superficie mayor que los cuerpos pequeños. De aquí que la Tierra sea hoy bastante activa, en tanto que los fuegos volcánicos de la Luna se agotaron hace mucho tiempo. (Los de Marte son probablemente activos hoy día, si bien con una intensidad muy inferior a los terrestres.) Io, que tiene el tamaño y la masa de la Luna terrestre, sólo habría de presentar, de acuerdo con lo anterior, signos de antiguo vulcanismo.

Esta expectativa cambió de signo en marzo de 1979, días antes de que el *Voyager 1* alcanzase el sistema joviano, cuando Stanton J. Peale, de la Universidad de California en Santa Bárbara y, Ray T. Reynolds y Patrick M. Cassen, del Centro de Investigación Ames, publicaron su análisis de las fuerzas de marea como proceso generador de calor para Io. Este y Europa están trabados por una de las resonancias gravitatorias estudiadas a principios del siglo XIX por Pierre Simon de Laplace. En cada una de tales resonancias, la interacción gravitatoria de los dos satélites ajusta sus distancias orbitales hasta que sus períodos orbitales son múltiplos comunes. (El período orbital de Europa es doble que el de Io.) De la resonancia entre Io y Europa se sigue que Europa está continuamente imponiendo cierta excentricidad, o discrepancia de la circularidad, a la órbita de Io. En virtud de ello, el punto subjoviano en Io (punto del cuerpo que queda más próximo a Júpiter) oscila, produciendo una marea o abultamiento que emigra hacia adelante y hacia atrás en el satélite y que varía de amplitud, calentando el cuerpo por rozamiento interno. Peale y sus colegas calcularon que la cantidad de calor podía ser sorprendentemente grande, por lo que predijeron la posibilidad de características volcánicas importantes.

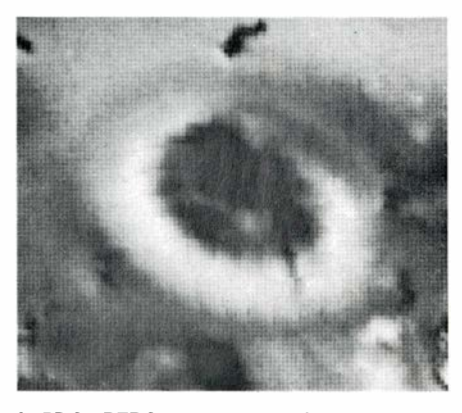
La opinión de Peale y sus colegas se fue asentando a medida que el *Voyager 1* se acercaba a Io. En imágenes de resolución cada vez mayor, no se hallaron trazas de cráteres de impacto en gran escala. Se divisaba una superficie cubierta de variopintas manchas rojas, anaranjadas y amarillas, que pronto se resolvieron en accidentes que se parecían a los ríos de lava y a las calderas (cráteres debidos a hundimiento volcánico) marcianos y terrestres. Io ofrecía una superficie geológicamente joven y activa. Pocos días después del paso del *Voyager 1*, un ulterior análisis de las imágenes realizado por los grupos científicos de navegación e imágenes del Voyager reveló erupciones volcánicas activas. Desde entonces, se ha venido investigando la naturaleza de la actividad volcánica recogida en las imágenes del Voyager y ha comenzado a dibujarse un esquema provisional de la cinética y termodinámica de dicho vulcanismo.

La mayoría de los rasgos que caracterizan a los volcanes de Io arrancan de la exótica química de su superficie. Bastante antes del *Voyager 1*, la ausencia de ciertas características en el espectro infrarrojo del satélite joviano (en particular, líneas de absorción debidas a hielo o escarcha) abogaban por una superficie extremadamente seca; en ciertos detalles de su espectro en longitudes de onda del visible se fundó la hipótesis de la existencia de azufre en alguna forma. Las líneas espectrales de emisión debidas a iones de azufre aprisionados en la magnetosfera de Júpiter eran también sugerentes. Varios instrumentos instalados a bordo de *Voyager 1* y *Voyager 2* revelaron la presencia de grandes cantidades de iones de azufre y oxígeno dispersos en la magnetosfera, si bien concentrados en la proximidad de la órbita de Io. Las emisiones ultravioleta de tales iones mostraron provenir de un toro situado alrededor de Júpiter y centrado en la órbita de Io. Finalmente, los miembros del grupo de espectrometría infrarroja del Voyager descubrieron características de absorción debidas al dióxido de azufre gaseoso sobre uno de los volcanes. Esto puso en claro que el azufre y sus compuestos constituyen el núcleo de la actividad volcánica de Io. Espoleados por los descubrimientos del Voyager, dos grupos de investigación, uno en el Laboratorio de Propulsión a Chorro del Instituto de Tecnología de California y otro en la Universidad de Hawai, identificaron un rasgo que destacaba en el espectro infrarrojo de Io cerca de la longitud de onda de 4,1 micrometros, procedente de escarcha de dióxido de azufre.

¿Por qué sobresale tanto el azufre en Io? Elemento común en el universo, viene a cifrarse su cantidad en la cuadragésima parte del contenido en oxígeno. El azufre escasea en la corteza terrestre, si bien se supone encerrado, en su mayor parte, en el núcleo en forma de sulfuro de hierro. También escasea en la Luna. Deficiencia que se inscribe en un problema mayor: el bajo contenido lunar de materias volátiles, que sigue constituyendo una de las principales dificultades que se oponen a la articulación de una hipótesis sobre el origen de dicho satélite. En los meteoritos, los primitivos sobre todo, abunda el azufre. Por lo que a la superficie de Marte concierne, los datos recogidos por los vehículos Viking que se posaron allí revelan una nutrida presencia de azufre en una forma u otra.

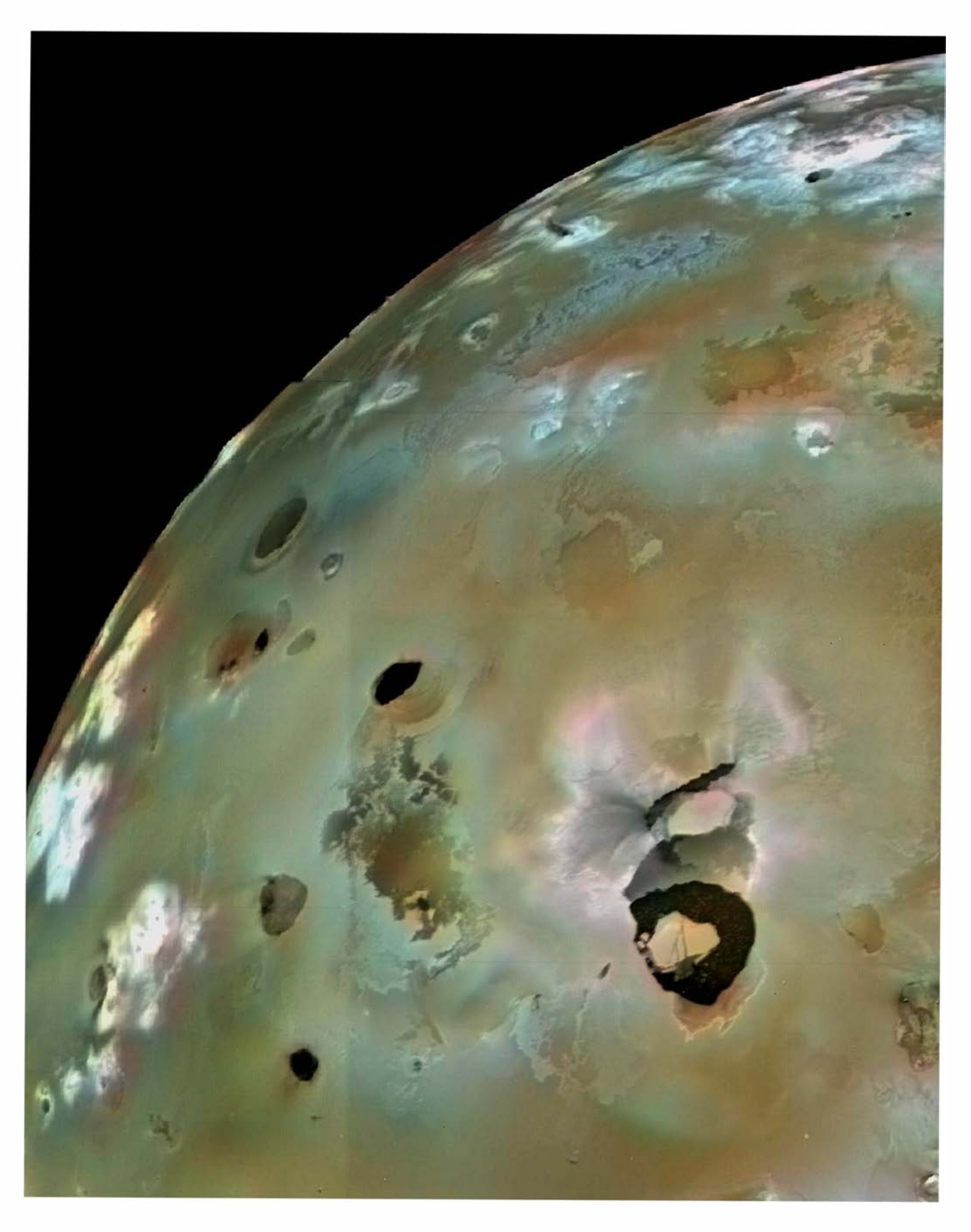






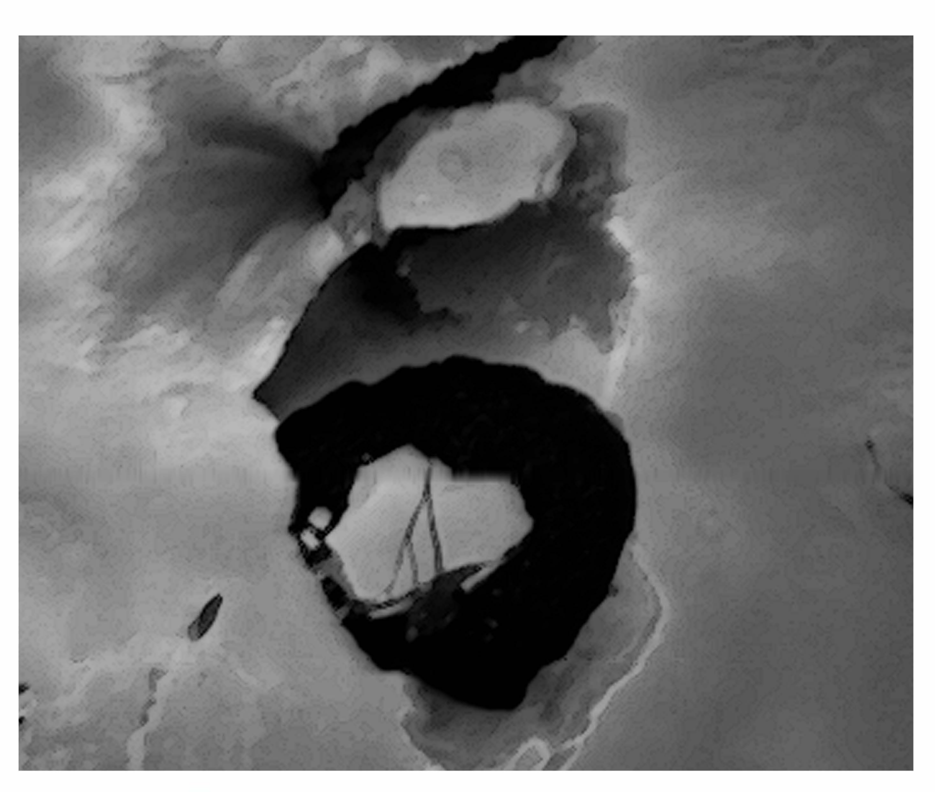
3. PROMETEO representa una clase de penachos volcánicos de Io menores, más fríos y de más larga duración que los semejantes a Pele. Estas tres vistas del penacho las tomó el *Voyager 1*. La superior nos lo muestra de perfil; mide 100 kilómetros de altura y 300 de ancho. Las otras dos tomas captan el penacho mientras el vehículo espacial pasaba por encima. Aunque el penacho aparece oscuro con respecto a la superficie, ha depositado un anillo de materia brillante. El brillo indica una probable acumulación de escarcha de dióxido de azufre.

La mayoría de los modelos relativos a la formación de los satélites de Júpiter sugieren que Io debió de ser, en su comienzo, más rico en sustancias volátiles que la Luna terrestre. Se ha de suponer entonces que el intenso calentamiento provocado por las mareas a lo largo de la historia del sistema solar expulsó del satélite todas las sustancias



4. PENACHOS LOKI: emergen de los extremos de una fisura lineal negra de unos 200 kilómetros de longitud. Cada uno ha producido un abanico brillante y un depósito mayor y oscuro, como si ambos penachos combinaran los estilos eruptivos de Prometeo y Pele. Por debajo de la fisura, hay una mancha negra en forma de *D*, que puede ser un lago de lava, aunque queda por establecer si la lava es azufre o silicato. Las medidas de la radiación infrarroja de Io realizadas por los vehículos espaciales Voyager indican que la mayor parte de la superficie del lago tiene una temperatura de unos 300 grados Kelvin, esto es, 170 grados superior a la que reina en el terreno circundante. La imagen es un mosaico del *Voyager 1* elaborado por la técnica de McEwen.

volátiles ligeras, incluyendo el agua; así pues, el azufre y sus distintos compuestos serían las únicas sustancias volátiles importantes que han quedado. Según una posibilidad rival, las condiciones en la vecindad del Io incipiente eran suficientemente cálidas para impedir la inclusión del agua, pero no hasta el punto de impedir la inclusión de azufre. La naturaleza de Amalthea, el satélite vecino a Io por la parte interna, podría decidir la cuestión: si el incipiente Io era demasiado cálido para el agua, Amalthea debería haber estado todavía más caliente (suponiendo que se formó en su actual lugar, con respecto a los demás satélites de Júpiter), y así se hallaría constituido por materia muy refractaria. Mas, a partir de datos espectrales limitados, se ha sugerido que la superficie de Amalthea se asemejaría a los meteoritos carbonáceos primitivos, lo que implica un importante contenido de materiales volátiles.



5. TOMA MAS PROXIMA DE LOKI, debida a la imagen de alta resolución hecha también por el *Voyager 1*. La gran almadía brillante en el lago está atravesada por diversas fisuras; varios trozos de materia brillante parecen desprendidos de sus bordes, moteando la superficie del lago. Tal vez la corteza del lago en enfriamiento se haya perforado por subversión convectiva o adición de nuevo material eruptivo. La balsa se supone constituida por azufre elemental. El lago mide 250 kilómetros de diámetro.

Hay que distinguir varios tipos de actividad volcánica en Io; todos ellos, determinados probablemente por la química del azufre. El tipo de los penachos eruptivos da cuenta de la más espectacular actividad volcánica observada hasta ahora en el satélite joviano. Fijémonos en Pele, que se encontraba en actividad durante el paso del *Voyager 1* y fue el primer penacho que se descubrió. Ascendió unos 300 kilómetros y distribuyó su materia eyectada en una suerte de paraguas de unos 1400 kilómetros de diámetro. La sombrilla, típica de muchos de los penachos, indica que la materia expulsada describe trayectorias balísticas. Por otro lado, un conjunto de imágenes del penacho de Pele registra una estructura arremolinada. Los remolinos sugieren turbulencia, y, consiguientemente, que las partículas están interaccionando con algún gas impulsor invisible. (Las imágenes del Voyager muestran, sobre todo, partículas sólidas de tamaño del micrometro, o inferiores; los gases asociados no pueden verse directamente.) Los modelos de simulación por ordenador de penachos formados por expulsión balística de partículas bajo la acción de la gravedad de Io reproducen muy bien las características generales de los penachos simétricos de dicho cuerpo celeste.

Cualquier modelo de los mecanismos responsables de la formación de los penachos tendrá que justificar las altas velocidades de expulsión deducidas del análisis balístico, que van desde 500 a 1000 metros por segundo. Bradford A. Smith, de la Universidad de Arizona, y sus colaboradores Eugene M. Shoemaker y Susan W. Kieffer, del Servicio de Inspección Geológica de los Estados Unidos, y Allan F. Cook II, del Centro de Astrofísica del Observatorio del Harvard College y del Observatorio Astrofísico Smithsoniano, han propuesto el modelo géyser; arrancan de la hipótesis según la cual la gran cantidad de calor que las mareas aportan a Io debe producir temperaturas bajo la superficie que permitan que el azufre líquido y el dióxido de azufre fundido se hallen en contacto a profundidades relativamente pequeñas. Kieffer ha realizado un estudio profundo de la termodinámica de las erupciones de penachos para una amplia variedad de condiciones aplicables a esas sustancias. En el extremo de baja energía dentro del dominio de posibles erupciones, hay una que comienza cuando el dióxido de azufre líquido entra en contacto con azufre fundido y caliente. El dióxido de azufre líquido comienza a hervir, y así empieza a expandirse una mezcla de líquido y gas elevándose hacia la superficie a través de algún conducto. A medida que continúa la dilatación, aumenta la producción de vapor. Entonces, el dióxido de azufre alcanza su punto triple, o temperatura a la cual coexisten el sólido congelado, el líquido y el gas. En estas circunstancias, y todavía a alguna distancia bajo la superficie de Io, el líquido restante se congela; la ulterior dilatación del dióxido de azufre viene señalada por la condensación de vapor de dióxido de azufre en una nevada de partículas sólidas. Una alternativa de mayor energía comienza con dióxido de azufre en estado de vapor sobrecalentado. En este caso, la dilatación hacia la superficie está acompañada primero por la condensación de líquido y después por la congelación y formación de nieve.

En cada caso, la mezcla de fases de dióxido de azufre es acelerada en el interior del conducto hasta que la corriente alcanza la velocidad del sonido (en ese medio). En este punto, su velocidad (cientos de metros por segundo) no puede ya aumentar. Más cerca de la superficie, el conducto puede ensancharse. Además, al alcanzar la superficie, la corriente se puede dilatar en el seno de la tenue atmósfera de Io. De aquí que la corriente llegue a alcanzar una impresionante velocidad inmediatamente encima del orificio de salida. Calculando la cantidad de energía disponible, Kieffer sostiene que son posibles velocidades de salida superiores a los 1000 metros por segundo. Pero ha



6. MAASAW PATERA y las marcas superficiales que rodean esa caldera de Io se parecen mucho a las que los geólogos ven en la Tierra y Marte. La principal diferencia concierne al tamaño: Maasaw Patera es grande, pues sólo la caldera mide 50 kilómetros de diámetro. Dos niveles en Maasaw Patera sugieren fases de colapso. La mayor parte de la caldera tiene 0,7 kilómetros de profundidad, según cálculos basados en las sombras de la estructura. La parte restante, en la porción superior derecha de la caldera, alcanza los dos kilómetros. En algunos puntos, el afloramiento de lava debió rebasar el borde de la caldera, produciendo corrientes de lava de cientos de kilómetros de longitud. La imagen es otro mosaico de McEwen con datos del *Voyager 1*.

señalado también que las observaciones del Voyager pueden explicarse admitiendo un amplio margen de condiciones en el subsuelo de Io; no es posible, pues, especificar la temperatura, profundidad del origen y composición de los distintos penachos.

La distribución de los penachos en la superficie y su comportamiento con el tiempo nos ofrecen ulterior información. Cuando llegó el *Voyager 1*, se vieron al menos nueve penachos en erupción; algunos se observaron varias veces los días que duró el paso. Los nueve, que se mantenían más o menos uniformemente distribuidos en longitud, se concentraban a latitudes bajas: ocho de ellos se encontraban en actividad cuatro meses más tarde, a la arribada del *Voyager 2*. (Pele, el penacho mayor, ya no estaba en actividad.) Había dos que formaban par: uno a cada extremo de un accidente en hendidura al Norte de una mancha negra llamada Loki. Presentaban una notable diversidad en la altura de su erupción durante los pasos y también entre ellos. La superficie que rodeaba a los orificios de Pele y Loki cambió bastante del paso de un Voyager al del otro.

Pele y Loki no distan mucho entre sí. Ambos están en un hemisferio intermedio entre el que se enfrenta a Júpiter y el que mira a la órbita de Io, desde atrás. (Io, como la Luna, presenta siempre la misma cara hacia el planeta.) Desde hace tiempo se sabe que, en promedio, ese hemisferio es mucho más rojo y oscuro que el opuesto. Los demás cambios importantes registrados en la superficie de Io de un paso a otro tuvieron lugar también allí. En torno a dos calderas se constituyeron sendos depósitos en forma de anillos, de un color rojizo oscuro; cada uno medía unos 1400 kilómetros de diámetro, el tamaño aproximado del depósito de Pele. Las calderas, sitas en latitudes altas, eran la llamada Surt, en el Norte, y Ayen Patera, en el Sur. Ninguna mostraba signos de erupción en el momento de pasar los Voyager; si bien la actividad en Surt pudo desencadenar un período de emisión térmica incrementada en la longitud de onda de cinco micrometros (en el infrarrojo), registrada en el Observatorio de Mauna Kea por William M. Sinton, de la Universidad de Hawai en Honolulu, un mes después del paso del *Voyager 2*. El depósito de Aten Patera se descubrió en un reciente análisis de las imágenes del proyecto Voyager por Alfred S. McEwen y uno de los autores (Soderblom), del Servicio de Inspección Geológica de los Estados Unidos.

Por su distribución de colores y tamaño, esos enormes depósitos recuerdan el que rodea a Pele; ante tal semejanza, McEwen y Soderblom sugieren que corresponden a una clase de erupciones grandes, breves y enérgicas, de las que Pele es el mejor ejemplo conocido. Sirviéndose de mediciones realizadas por el espectrómetro de infrarojos del *Voyager 1*, John C. Pearl y sus colegas, del Centro de Vuelos Espaciales Goddard, de la NASA, estiman que la temperatura del orificio de Pele es de unos 650 grados Kelvin; las observaciones de Sinton, presumiblemente de Surt, apoyan una cifra similar.

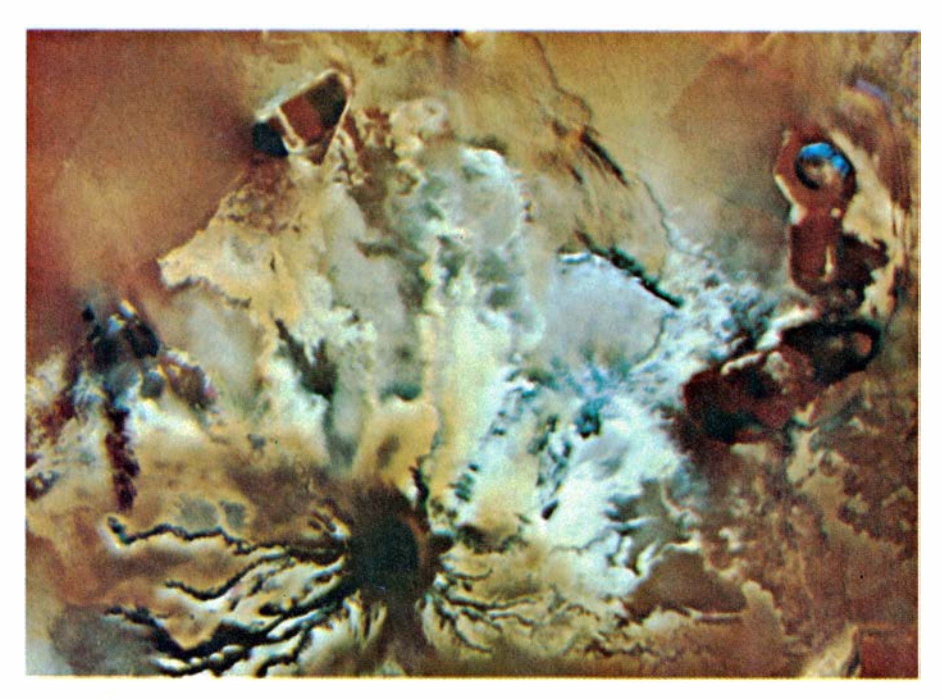
Las erupciones de una segunda clase son más pequeñas: depositan anillos de unos 300 kilómetros de diámetro. Persisten largo tiempo. Al haberlas contemplado ambos Voyager, se les atribuye una vida activa de varios años, por lo menos. Parecen ser más frías: alrededor de 400 grados Kelvin. Están limitadas a una banda ecuatorial de materia blanca brillante, que se supone rica en escarcha de dióxido de azufre. (Sus depósitos en forma de anillo están constituidos por la misma materia.) El mejor ejemplo nos lo ofrece Prometeo, un accidente a cinco grados al Sur del ecuador. Diríase que el par de penachos de la fisura de Loki oscila entre las dos clases. Los rodea un depósito complicado: cada uno tiene una zona interior que consta de un depósito en forma de abanico, probablemente rico en dióxido de azufre, de unos 200 kilómetros de largo, y una zona externa semejante en tamaño y color a los depósitos que ciñen los tres penachos grandes y análogos al de Pele.

En opinión de McEwen y Soderblom, las dos clases de penachos corresponderían a dos sistemas volcánicos claramente diferentes, impulsados por distintos cuerpos volátiles. Y siguen: esas clases provienen de la peculiar relación que se da en el azufre entre temperatura y viscosidad. Cuando éste se calienta más allá de su punto de fusión, experimenta cambios secuenciales. Empieza por sólido amarillo; después, a 400 grados Kelvin de temperatura, se funde y constituye un fluido amarillo de baja viscosidad. Un calentamiento ulterior trastorna su color, primero naranja y después, a 430 grados, rosa claro. Se vuelve luego rojo oscuro y se espesa; hacia los 500 grados, se transforma en una especie de alquitrán negruzco. A los 600 grados, su viscosidad decrece y, a 650 grados, se torna fluido de nuevo. Por último, a una temperatura que depende del proceso, se vaporiza.

El azufre fundido de baja viscosidad, que puede transportar calor fácilmente, se da en dos formas: una rojiza clara, a temperaturas entre 400 y 430 grados Kelvin, y otra negra opaca, a temperaturas superiores a 650 grados. La idea, por tanto, es que los pequeños penachos, así Prometeo, surgen impulsados por el azufre líquido rojo al entrar en contacto con dióxido de azufre líquido, provocando la vaporización del dióxido de azufre y su dilatación, según proponen Smith, Shoemaker, Kieffer y Cook, mientras que los grandes penachos, Pele por ejemplo, se producen cuando los silicatos calientes de la corteza de Io vaporizan azufre a temperaturas entre 700 y 1200 grados Kelvin.

La notable diferencia en la duración de las dos clases de penachos se explica mediante este modelo. Los pequeños penachos de dióxido de azufre persisten años enteros porque el dióxido de azufre se mantiene líquido en un amplio intervalo de profundidades en la corteza de Io. El líquido tiene una viscosidad extremadamente baja (que viene a coincidir con la del alcohol a temperatura ambiente); se desliza fácilmente a través de la corteza para alcanzar las salidas en la superficie. Por contra, los grandes penachos de azufre negro duran sólo unos días, pues las condiciones que favorecen sus erupciones están delicadamente equilibradas. La expulsión de azufre podría detenerse, por ejemplo, si la temperatura bajase mucho de los 650 grados en el orificio, en el conducto o a mayor profundidad en la corteza, donde se produce el encuentro del azufre y los silicatos, ya que el azufre se enfriaría rápidamente adoptando la forma de sustancia bituminosa y casi sólida.

Queda sin resolver la cuestión de por qué ambas clases de penachos están distribuidas de la manera que lo están. Las imágenes de los Voyager del hemisferio rojo oscuro de Io, donde se han encontrado los grandes penachos, muestran numerosas montañas que deben estar constituidas por silicatos que asoman por encima de las llanuras de Io. (Una montaña de azufre no sería capaz de soportar su propio peso.) Las imágenes del hemisferio opuesto tienen una resolución más baja. En algunas de ellas, sin embargo, la región donde más abundan los penachos pequeños, del tipo Prometeo, es en el terminador (límite del día y la noche), en que las



7. RA PATERA y las marcas de la superficie circundante difieren notablemente de lo que los geólogos ven en la Tierra y en Marte. Aquí, Ra Patera está en la parte inferior de la imagen. Nos hallamos ante una caldera llena de materia negra que emite largos y sinuosos ríos, principalmente hacia la izquierda, hasta 200 kilómetros de distancia. David C. Pieri, del Laboratorio de Propulsión a Chorro del Cal Tech, y Carl Sagan, de Cornell, proponen que el centro oscuro de Ra Patera representa una fase bituminosa y negruzca del azufre, y que los ríos constituyen fases más frías y menos viscosas. Hay dos calderas notables más. Hacia la parte superior izquierda se distingue una con bandas amarillas, anaranjadas y negras. Los colores pueden obedecer a distintas fases del azufre. En la parte superior derecha se aprecia una caldera circular que incorpora un cuarto creciente azulado. En una imagen tomada por el *Voyager 1* seis horas antes de la representada, el creciente no aparece. Tal vez el creciente indica una erupción de baja energía.

montañas deberían ser fácilmente visibles. En otras imágenes se localiza en el borde iluminado de Io. Estas imágenes muestran principalmente llanuras lisas.

De acuerdo con ello, McEwan y Soderblom abogan por una morfología muy variada de la corteza de Io. En particular la corteza de la región de los penachos grandes, de tipo Pele, será mucho más delgada que en los demás lugares, de suerte que los silicatos calientes estén más próximos a la superficie y la presión que confina al vapor de azufre sea menor. Además, la corteza en la región de los pequeños penachos, del tipo Prometeo, puede estar enriquecida en dióxido de azufre líquido en virtud de algún proceso geofísico (un gradiente de presión hidrostático que variase con la latitud como consecuencia de la rotación de Io), que impulse el líquido hacia el ecuador. Vale la pena señalar que el par de penachos Loki, que constituyen quizás un híbrido entre las dos clases de penachos bien diferenciadas, caen en el área de superposición de las regiones en donde se encuentran las dos clases.

Vayamos a otra clase de actividad volcánica, muy diferente de los penachos de Io semejantes a géyseres. Denuncian su presencia las grandes calderas volcánicas y sus ríos de "lava" asociados y otras marcas en la superficie, que se encuentran entre los accidentes más patentes identificados en las imágenes Voyager. En general, las calderas se parecen a las terrestres, aunque mayores: dos centenares largos de ellas tienen diámetros de 20 kilómetros por lo menos; en la Tierra, con una superficie sólida tres veces y media mayor, sólo hay 15 de esos. A diferencia de los penachos, las calderas de Io no muestran concentración relacionada con la latitud, y se reparten, igual por unidad de superficie, en altas latitudes y en la vecindad del ecuador.

El debate acerca de la naturaleza de las calderas se ha centrado en las cuestiones planteadas por los penachos, es decir, las condiciones térmicas dominantes bajo la superficie y la disponibilidad de azufre elemental y dióxido de azufre. Dos modelos antagónicos sobre el particular serían el del azufre y el de los silicatos. De acuerdo con el primero, propuesto en su forma más pura por Carl Sagan, de la Universidad de Cornell, las lagunas subsuperficiales de azufre fundido, postuladas por Smith y sus colaboradores para explicar los penachos de dióxido de azufre en Io, también darían cuenta de los ríos volcánicos de la superficie; los ríos que rodean las calderas serían el resultado de "lavas" de azufre líquido, en lugar de basalto fundido y otros silicatos. De acuerdo con el segundo, propuesto por Harold Masursky y Michael H. Carr, de la Inspección Geológica, y sus colaboradores, las calderas representarían un vulcanismo de silicatos muy parecido al de la Tierra. Con una sola diferencia: los ríos de silicatos en Io están teñidos de azufre.

El modelo de azufre está avalado, en Io, por la innegable abundancia de este elemento y de sus compuestos, por la detección de gas dióxido de azufre en la vecindad de una de las erupciones y por la probabilidad de que esas sustancias alcancen sus puntos de fusión a profundidades no muy grandes en la corteza. Se ha sugerido, además, que la secuencia de cambios de color y de viscosidad del azufre al calentarse o enfriarse explicaría el cambio de color con la distancia observado alrededor de muchas calderas. Ra Patera es un buen ejemplo. David C. Pieri, del Laboratorio de Propulsión a Chorro, y Sagan proponen que las oscuras regiones centrales de Ra Patera representan la fase negra de alta temperatura del azufre fundido. Los largos ríos digitiformes que irradian desde la caldera indican corrientes más frías y menos viscosas de las fases roja y anaranjada.

Pero esta hipótesis encuentra una dificultad en el comportamiento de las diversas fases de azufre bajo las condiciones de Io; lo desconocemos. Es verdad que los colores de las fases de alta temperatura se pueden retener en el laboratorio enfriando rápidamente el líquido. Sin embargo, las fases sólo son metastables y revierten al azufre sólido ordinario transcurrido algún tiempo. Además, es probable que las diversas fases tengan impurezas, lo que complicaría más la cuestión.

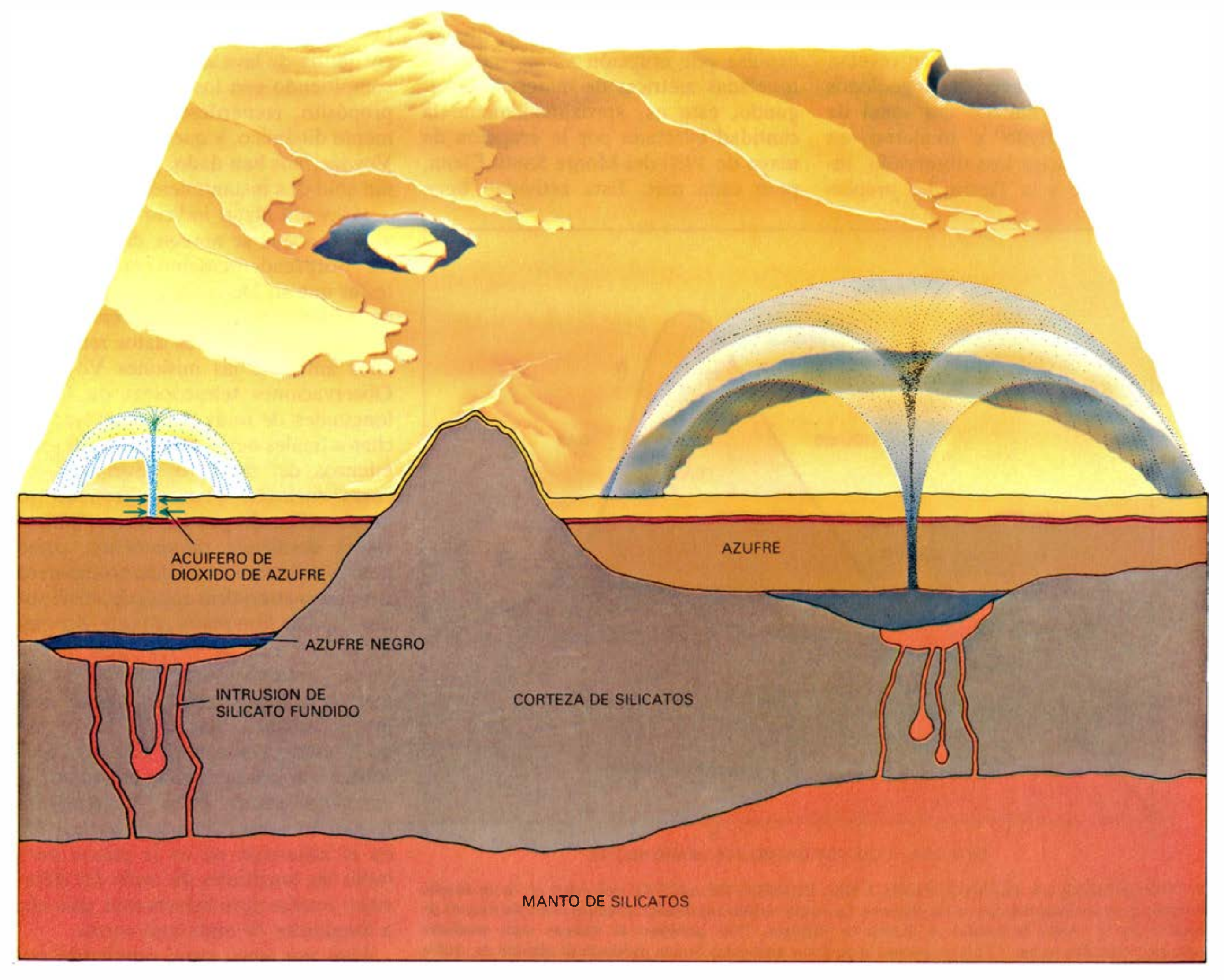
La defensa del vulcanismo de silicatos se apoya, en parte, en la topografía medida de Io. No es fácil tomar aquí medidas: las variaciones espaciales de brillo en las imágenes Voyager, resultantes de diferencias en la composición de la superficie, tienden a ser mucho más fuertes que las variaciones en brillo resultantes del relieve topográfico. Pese a ello, se ha intentado sacar algunas en las imágenes de alta resolución obtenidas bajo iluminación favorable durante el paso más próximo a Io, el del *Voyager 1*. Las medidas sugieren una variabilidad topográfica mucho mayor que la de los otros satélites galileanos: sus montañas tienen alturas de cinco a diez kilómetros, y al menos algunas de sus calderas se hunden en pro-

fundidades de dos a tres kilómetros. El mantenimiento de una tal topografía sería ciertamente difícil en un cuerpo cuya corteza incluyese extensos depósitos de azufre caliente, posiblemente líquido, y de dióxido de azufre. Una corteza más resistente cuyo esqueleto, al menos, fuese de silicatos parece necesaria; en consecuencia, la amplia distribución de calderas en la superficie excluiría la existencia de extensos "océanos" subsuperficiales de azufre.

El otro argumento principal esgrimido por los proponentes del vulcanismo de silicatos sostiene que las estructuras que se aprecian en Io (calderas, largos ríos semejantes a dedos, montículos volcánicos con cráteres centrales, etcétera) coinciden, a grandes rasgos, con las que los geólogos están acostumbrados a ver en la Tierra y en Marte. No parece plausible que el azufre, materia volcánica poco corriente y con una rara relación entre temperatura y viscosidad, reprodujera tan de cerca las estructuras creadas por las lavas basálticas. Desgraciadamente, los geólogos sólo tienen una limitada experiencia con el azufre como fluido volcánico. Hay algunos raros ríos de azufre en la Tierra, pero se dan sólo allí donde el vulcanismo de silicatos funde depósitos de azufre.

¿Qué está, pues, ocurriendo en Io? Cada bando del debate (azufre contra silicato) concede ahora la posibilidad de que el vulcanismo propuesto por su opositor se produzca también en algunos lugares de Io, al menos. Gerald G. Schaber, de la Inspección Geológica, observa, por ejemplo, que ciertas calderas están estampadas en regiones montañosas de Io. Puesto que cabe presumir que las montañas sean de silicatos, las calderas lo serán también. Ambos bandos convienen, además, en que la actividad volcánica en la superficie de Io, cualquiera que sea su naturaleza, está impulsada por silicatos fundidos en las profundidades de la corteza. La cuestión a resolver es si uno de los dos tipos de actividad volcánica, azufre o silicato, domina en la superficie.

Queda por describir un rasgo volcánico más. Se trata de una mancha negra al sur de los penachos Loki; ¿acaso un gran lago de lava de la clase que típicamente llena las calderas activas durante las erupciones en la Tierra? En la época del paso del *Voyager 1*, este accidente era la mayor "mancha caliente" en Io: tenía una temperatura de unos 300 gra-

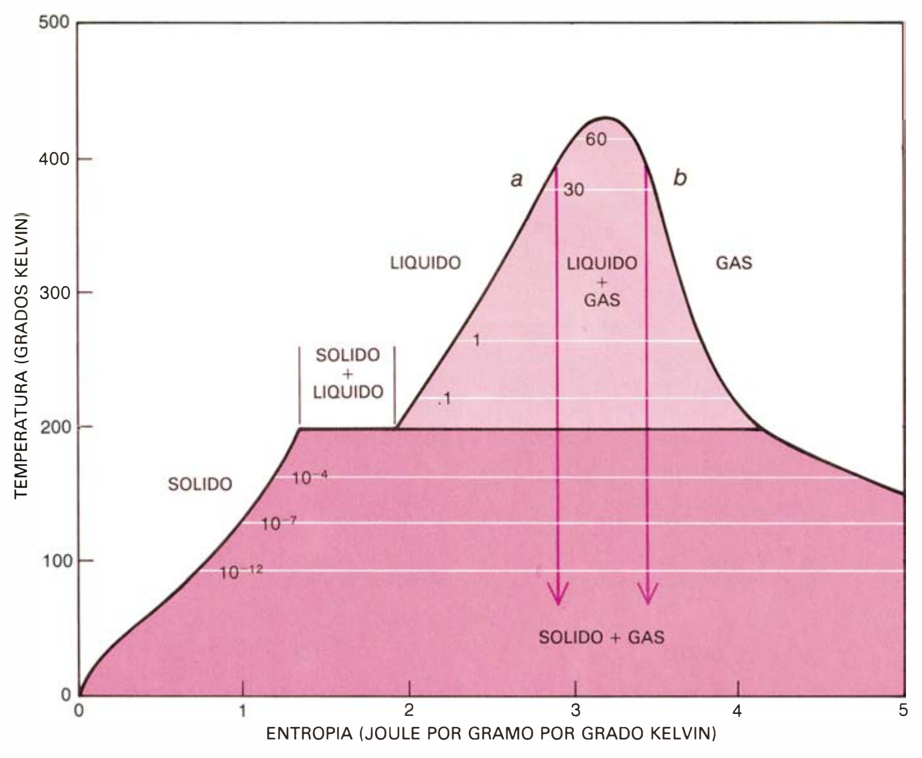


8. CORTE VERTICAL HIPOTETICO DE IO que resume una explicación posible de la actividad volcánica del cuerpo. En la sección, el manto es de silicato fundido en todo o en parte (*anaranjado*); sobre la corteza, de silicato sólido (*castaño*), yacen capas de residuos de silicatos ricos en azufre. El azufre de la parte superior es frío y sólido (*amarillo*); después viene una delgada zona en la que el azufre está fundido (*rojo*) y una zona más espesa donde el azufre es caliente y bituminoso (*castaño amarillento*). En la parte izquierda se ha modelado un penacho del tipo Prometeo. El azufre fundido entra en contacto con el dióxido de azufre (*azul*) líquido, que asciende hirviendo por un canal, produciendo, ya en la superficie, un penacho rico en nieve de dióxido de azufre. A la derecha se ha modelado un penacho como Pele. Aquí, la intrusión de silicato calienta el azufre; éste adopta una fase muy caliente y bituminosa (*negro*). La verdad es que se vaporiza, y, vaporizado, hace erupción. Se dibuja un río de lava como Loki en la parte superior izquierda; en la parte superior derecha, una caldera. Su fondo es la corteza de silicato subyacente a una delgada capa de azufre. (Recreación gráfica realizada por Ian Worpole.)

dos Kelvin, mientras que la temperatura local de fondo alcanzaba sólo los 130 grados. Además, las imágenes de alta resolución mostraron que el lago contenía una "almadía" de materia brillante. La almadía parecía estar atravesada por grietas, y pequeños trozos del mismo material daban la impresión de haberse desprendido de sus bordes. Diríase que la corteza del lago, cuando estaba enfriándose, se hubiese roto por subversión convectiva o por adición de nueva materia eruptiva. El accidente es bastante mayor que los lagos de lava de las calderas hawaianas. Su longitud de unos 250 kilómetros abarcaría la totalidad de la cadena de las islas hawaianas. ¿Está llena de silicato fundido o de azufre fundido que se ha enfriado ahora? No lo sabemos.

La prueba más clamorosa de una superficie joven es la total ausencia de cráteres de impacto, según revelan las imágenes Voyager. Los geólogos planetarios han hallado esa señal de bombardeo antiguo y moderno en cuantas superficies han observado, incluidos Marte y la Tierra; las propias superficies picadas de cráteres de Ganimedes y Calixto muestran que los satélites jovianos han sufrido una historia de impactos similar a la de los planetas interiores. Además, cálculos realizados por Shoemaker sugieren que el actual flujo de cometas y asteroides a través del sistema joviano debería estar produciendo cráteres grandes en Io a una velocidad dentro de un orden de magnitud de aquella a la cual se están formando en la actualidad cráteres en la Luna.

A partir del descubrimiento de Shoemaker, hemos calculado la actividad volcánica que se requiere para recubrir los cráteres de Io. Llegamos a la conclusión de que a lo largo de un año, en promedio, la superficie del satélite debe recubrirse de una capa de detritus volcánicos de un milímetro de espesor. Esto implica un cuerpo realmente muy activo. Implica, en particular, que Io expulsa por erupción varios miles de toneladas métricas de materia por segundo, esto es, aproximadamente la cantidad eyectada por la erupción de mayo de 1980 del Monte Santa Elena, pero cada mes. Esta actividad tiene consecuencias importantes. Entre otras cosas, sugiere que al menos la parte superior del manto y corteza de Io se ha reciclado muchas veces en el curso de la historia del satélite. Este reciclado concuerda con la idea según la cual el calentamiento por fuerzas de marea ha "hecho sudar" a Io, hasta obligarle a perder las materias volátiles que en alguna época hubiera tenido.

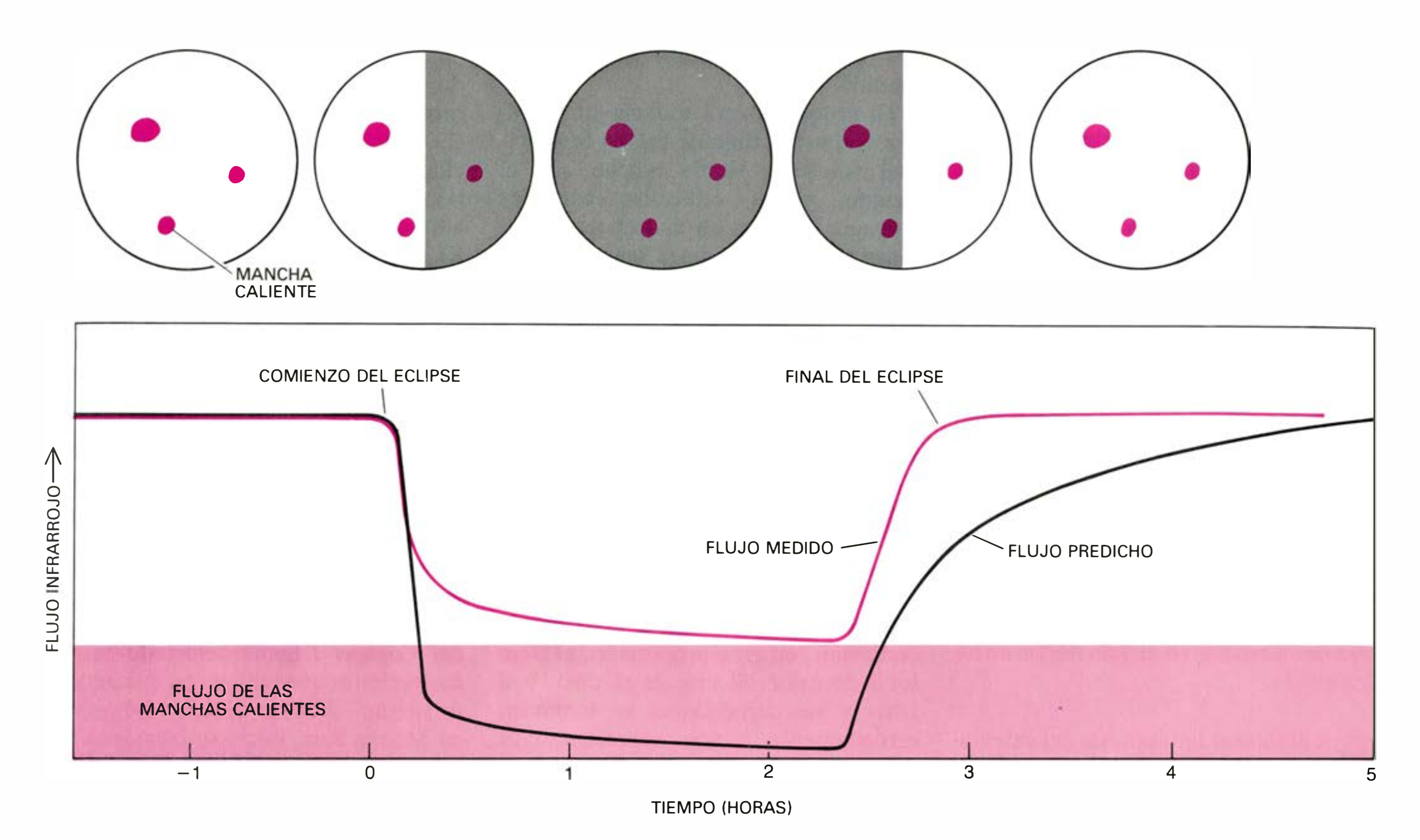


9. NOS SUGIERE LA TERMODINAMICA DEL DIOXIDO DE AZUFRE que existe en Io un amplio dominio de erupciones del tipo de los géyseres. La escala vertical representa la temperatura del dióxido de azufre; en la escala horizontal se ilustra la entropía. (Las presiones se indican aquí mediante isobaras rotuladas en bar.) En el esquema, la posición horizontal de una partícula de dióxido de azufre señala proporciones de fases (por ejemplo, líquido y gas). Además, una línea vertical traza la evolución seguida por una erupción de dióxido de azufre que no realiza trabajo (tal como empujar contra una barrera) y no gana ni pierde calor. Se representan dos erupciones de esas. Una comienza cuando el dióxido de azufre líquido se calienta hasta 393 grados Kelvin, temperatura del azufre fundido en las profundidades de la corteza de Io (*a*). El dióxido de azufre hierve así elevándose hacia la superficie. La otra erupción (*b*) comienza con dióxido de azufre gaseoso. En ambos casos, el dióxido de azufre explota formando una nube de hielo y gas en el casi vacío que reina en la superficie de Io. La termodinámica de estas evoluciones fue analizada por Susan W. Kieffer, del Servicio de Inspección Geológica americano.

No se conocen bien los pormenores del reciclado. Por ejemplo, no está claro cuán importantes sean los penachos y los ríos de lava en cuanto a recubrir de materia nueva la superficie. Los cálculos de la cantidad total de materia expulsada en los penachos de Io indican que la velocidad de recubrimiento atribuible solamente a penachos varía de algunas diezmilésimas de milímetro por año a una décima de milímetro o más. Lo que sugiere que las corrientes de lava están, por lo menos, compitiendo con los penachos. A este propósito, recuérdese que Io es altamente dinámico, y que los pasos de los Voyager nos han dado, por decirlo así, tan sólo dos instantáneas.

¿Hay allí algún indicio de actividad volcánica que se aprecie desde la Tierra? Sorprendentemente hemos de contestar que sí. De hecho, había pruebas de la existencia de volcanes encerradas en ciertos enigmáticos datos recogidos años antes de las misiones Voyager. Observaciones telescópicas de Io en longitudes de onda del infrarrojo, hechas a finales del decenio de 1960 y comienzos del de 1970, revelaron dos cosas curiosas. Primero, cuando Io queda eclipsado por Júpiter, su temperatura desciende rápidamente, como podría esperarse de un cuerpo con poca o ninguna atmósfera cuya superficie sea una capa de fino polvo aislante. Sin embargo, al revés de lo que pasa con la Luna, la temperatura mínima en Io, que se alcanza cuando el satélite está profundamente inmerso en la sombra de Júpiter, es demasiado alta para concordar con la hipótesis de una capa aislante homogénea. Además, la temperatura de Io deducida a partir de su brillo en el infrarrojo no es la misma para todas las longitudes de onda del infrarrojo; resulta apreciablemente más alta a longitudes de onda más cortas.

Uno por uno, estos problemas podían abordarse, aunque con hipótesis ad hoc. Se idearon modelos de dos capas de Io para explicar los datos de los eclipses; se propusieron variaciones de la emisividad dependientes de la longitud de onda para justificar las ano-



10. ENFRIAMIENTO DE IO durante un eclipse. Intrigó a los observadores que midieron el flujo infrarrojo de Io al pasar éste por la sombra de Júpiter. Creían que el flujo descendería rápidamente hasta cero al comienzo de cada eclipse y retornaría a su valor normal cuando la luz solar calentase la superficie (*curva negra*). No fue así. El flujo permaneció por encima de cero (*curva en color*). La explicación que se da es que las "manchas calientes" volcánicas emiten un flujo de calor más o menos constante durante todo el eclipse. Ello significa que la actividad volcánica de Io puede vigilarse desde la Tierra.

malías de temperatura. Después de los pasos de los Voyager y del descubrimiento de volcanes, tres investigadores del Laboratorio de Propulsión a Chorro, Dennis L. Matson, Gary A. Ransford y uno de los autores (Johnson), retornaron a los enigmáticos datos, esperando resolver los viejos problemas a la luz de la nueva información. Idearon un sencillo modelo en el que la superficie aislante de Io quedaba moteada de manchas calientes a determinada temperatura. Los problemas desaparecían si se admitía que el uno por ciento de la superficie, aproximadamente, se hallaba cubierto por las manchas calientes cuyas temperaturas fueran semejantes a las que los vehículos espaciales Voyager midieron.

El modelado proporcionó un valor de la cantidad de energía radiada por las manchas calientes: promediada sobre la superficie de Io, valía unos dos watt por metro cuadrado. Nuevos análisis de los datos Voyager y nuevas observaciones infrarrojas realizadas desde Mauna Kea produjeron estimas independientes de la cantidad; concordantes con el modelo y los datos primitivos. La mejor estima actual del flujo medio de calor en Io, inspirada en todas esas fuentes, es de 1,5 watt por metro cuadrado más o menos 0,5 watt. En comparación, la Tierra viene a radiar 0,08 watt por metro cuadrado, y la Luna, con casi el mismo tamaño y masa que Io, sólo 0,03. La radiación de Io destaca por lo anómala. Sin apelar siquiera al calentamiento por marea, debe existir una fuente de energía distinta de la radiactividad interna.

Hemos de señalar que el flujo de calor en Io difiere del flujo de calor en la Tierra y en la Luna, cuantitativa y cualitativamente. En la Tierra y en la Luna, casi todo el exceso de calor interno generado por radiactividad alcanza la superficie por conducción a través de la corteza. (Las erupciones volcánicas de la Tierra sólo dan cuenta de una pequeña parte del total.) Desde allí se radia al espacio sin aumentar de manera apreciable la temperatura del cuerpo. (La superficie se calienta gracias, sobre todo, a la energía que recibe del Sol.) Esto significa que el exceso no se puede medir a cierta distancia como emisión infrarroja. Hace falta una medida precisa del gradiente térmico en la parte superior de la corteza. La situación en Io es distinta. El exceso de calor interno alcanza la superficie, principalmente, por convección: es transportado por fluidos calientes que se elevan hasta las manchas calientes. Desde allí, el calor se radia al espacio a temperaturas mucho más altas que la temperatura superficial del satélite; se puede, pues, medir sin mayor dificultad. La energía emitida por la superficie de Io está entre 10^{13} y 10^{14} watt.

Como fuentes de esta energía se han propuesto dos procesos, distintos de la radiactividad interna: el calentamiento producido por las mareas y el calentamiento producido por la resistencia del satélite a la corriente eléctrica inducida en el cuerpo en virtud de su interacción con la magnetosfera de Júpiter. Thomas Gold, de Cornell, sugiere que los penachos de Io proceden de tal interacción cuando la materia ionizada asciende a lo largo de las líneas de corriente del campo fluyendo hasta la superficie en forma de grandes chispas eléctricas sostenidas. Basa su tesis en la dificultad de impulsar gases volcánicos a velocidades que excedan la del sonido y en datos de los magnetómetros de las misiones Voyager, según los cuales en la vecindad de Io fluye una corriente de alrededor de un millón de ampère a lo largo de las líneas del campo magnético. (Una corriente de un millón de ampère es también lo que los cálculos teó-

ricos del flujo de iones en la magnetosfera joviana permitían esperar.)

La emisión de energía de Io, de 10^{14} watt, puede aprovecharse para poner a prueba la presumible importancia de los efectos electromagnéticos en el balance global de energía del objeto. Supóngase que toda la corriente que incide sobre Io se convierte en calor en el interior por calentamiento resistivo. La cantidad neta de calor variaría de 10^{11} a 10^{12} watt, no más de la centésima parte de la cantidad necesaria para explicar la emisión de calor. Ahora bien, los efectos electromagnéticos podrían influir en determinados episodios: los penachos eruptivos, por ejemplo. En tal caso, el "rayo" que produjese un penacho daría lugar a una mancha en Io con una temperatura efectiva del orden de 100.000 grados Kelvin. La búsqueda de tales manchas en el lado nocturno ha fracasado.

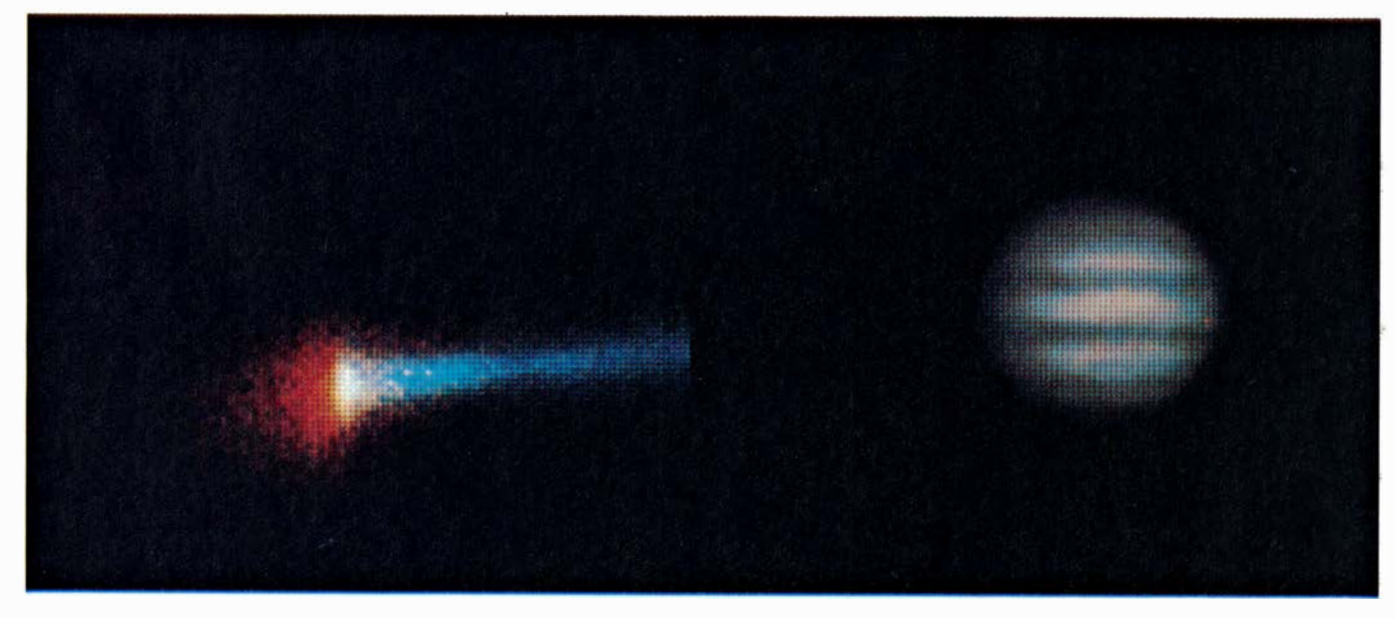
En el ínterin la hipótesis del calentamiento por mareas conocía un nuevo problema. La interacción mareal entre Io y Júpiter provoca abultamientos de marea en ambos cuerpos; el abultamiento en Júpiter aplica un par gravitatorio a Io, haciendo que el satélite se acelere y cambie a una órbita todavía más alta. (Una interacción semejante es responsable del lento y continuo alejamiento de la Luna con respecto a la Tierra.) Las mejores estimas actuales de la intensidad de la interacción marcan un límite superior a la energía media que llega a Io, y ese límite superior es aproximadamente la mitad del flujo de calor desde Io que se ha medido.

El problema está todavía sin resolver. Tal vez el flujo de calor y la actividad volcánica varíen mucho con el tiempo, y los vehículos espaciales Voyager llegaron en una época de intensa actividad. O quizá sean erróneas algunas de las variables e hipótesis involucradas en los cálculos de la evolución orbital de Io y del calentamiento por mareas. Y cabe también que las manchas calientes no guarden una distribución uniforme en la superficie de Io. Datos recientes recogidos en el Centro del Telescopio Infrarrojo de la NASA en Hawai por investigadores del Laboratorio de Propulsión a Chorro y la Universidad de Hawai, entre ellos uno de los autores (Johnson), sugieren que la región en torno a Loki puede contribuir, en gran proporción, al flujo total de calor. Si éste es el caso, y si Loki y sus alrededores se tomaran, erróneamente, como representativos del cuerpo entero, la energía total emitida por Io estaría más próxima al valor teórico obtenido.

No parece haber indicios de que el flujo de calor varíe mucho; los datos recogidos en un eclipse a principios del decenio de 1970 ofrecen resultados concordantes con las medidas de los Voyager y con datos de otro eclipse registrados en los últimos años. Hay un tipo de emisión infrarroja que exhibe espectaculares variaciones a corto plazo: la emisión cerca de la longitud de onda de cinco micrometros muestra brotes ocasionales, relacionados probablemente con procesos en Io variables de un día a otro. El primero de tales brotes fue detectado en 1978 por Fred C. Witteborn, del Centro de Investigación Ames, y sus colaboradores a bordo del avión instrumentado de gran altitud llamado Observatorio Volante Kuiper. Su breve observación de un flujo, en cinco micrometros, de más del doble de la intensidad que podría atribuirse a luz solar cabe calificarla de sorprendente.

El flujo sugería que una pequeña superficie, la diezmilésima del área de Io, mostraba una temperatura de unos 600 grados Kelvin en un momento en que la actividad volcánica parecía punto menos que impensable. Tras el paso del *Voyager 1*, la importancia de la observación de Witteborn se apreció de inmediato: el espectrómetro infrarrojo del *Voyager 1* había detectado pequeñas regiones que tenían una temperatura similar. William Sinton, trabajando en Mauna Kea, inició su búsqueda sistemática de más "sucesos de cinco micrometros". Uno de sus primeros éxitos fue la detección del suceso entre los pasos de los Voyager: el suceso que se cree ligado al cambio alrededor de la caldera de Surt.

Desde entonces, Sinton ha venido acumulando un valioso archivo de datos registrados a lo largo de varios años. Pocos episodios rivalizan con los primeros, aunque se advierte una continua variación en el flujo desde las "manchas calientes", que lo son mucho. La cantidad total de energía



11. TORO DE AZUFRE IONIZADO que circunda Júpiter y está centrado en la órbita de Io. Deriva de la actividad volcánica del satélite, que, según se cree, expulsa alrededor de una tonelada métrica de átomos de azufre y oxígeno por segundo hacia el espacio circundante a Júpiter. La imagen de un borde del toro fue realizada en los observatorios del Monte Wilson y Las Campanas por John T. Trauger, del Cal Tech. El falso color representa la proporción de átomos de azufre ionizado una sola vez (*azul*) y átomos doblemente ionizados (*rojo*). El borde derecho aparente marca el límite de los datos de Trauger.

que radian esas pequeñas zonas de alta temperatura es mucho menor que la cantidad que emiten zonas más extensas a menores temperaturas. De aquí que los sucesos a cinco micrometros no contribuyan, en cuantía apreciable, al balance energético global de Io. Sin embargo, proporcionan alguna luz sobre los violentos sucesos que presumiblemente generan penachos grandes del tipo Pele. Y es de esperar que faciliten la búsqueda de indicaciones de lavas de silicatos. En una perspectiva más amplia, las observaciones de Io realizadas desde la Tierra, incluidos estudios de eclipses y medidas en el infrarojo, constituyen medios potenciales para vigilar la actividad volcánica; proporcionarán un importante enlace entre los datos de los Voyager y los que se espera acumular cuando se envíe el próximo vehículo al sistema joviano.

Ese próximo vehículo será *Galileo*. Se prevé su lanzamiento para 1986 y su llegada a Júpiter en agosto de 1988; llevará una sonda que penetrará en la atmósfera de Júpiter y un vehículo orbital de larga vida que permitirá obtener unos 20 meses de observaciones detalladas del planeta, su magnetosfera y sus satélites. En su camino hacia Júpiter, *Galileo* pasará a 1000 kilómetros de Io, 20 veces más cerca que el paso más próximo de los Voyager. Los instrumentos a bordo de *Galileo* cartografiarán la superficie de Io y construirán imágenes a una resolución comparable a las de las imágenes landsat de la Tierra. Podrá examinarse la interacción entre la magnetosfera de Júpiter y la tenue atmósfera de Io; la radiotelemetría del vehículo espacial impondrá nuevas condiciones a los modelos del interior de Io.

Tras ese paso inmediato, *Galileo* evitará la vecindidad de Io. De no hacerlo, una prolongada dosis de radiación lo pondría fuera de servicio. Pese a ello, la mayor parte de la misión de 20 meses del vehículo orbital lo mantendrá dentro de la distancia desde la que los Voyager detectaron los volcanes. En una docena de ocasiones, *Galileo* explorará Io en busca de cambios volcánicos, construirá mapas térmicos que muestren la situación de las manchas calientes y buscará pruebas de iones que crucen de Io a la magnetosfera de Júpiter. Io habían estado intrigando a los astrónomos durante 400 años, antes de que los Voyager revelaran un mundo más complejo y fascinante. Y en ese camino avanzará la próxima fase de exploraciones.

Una comunidad agrícola de la Edad del Hierro en Europa central

Las excavaciones realizadas en un yacimiento de Baviera revelan un sistema económico que, desarrollado desde el 1000 hasta el 800 a. de C., permitió a los agricultores trocar sus excedentes con productos de artesanos especializados

Peter S. Wells

Aunque el modo de vida urbano era ya algo común en el Próximo Oriente hacia el 3500 a. de C., la mayoría de los poblados europeos situados al norte de los Alpes constituyeron pequeñas aldeas hasta fines de la Edad del Bronce e inicios de la Primera Edad del Hierro. En esa región, la transición al modo de vida urbano ocurrió hacia el 800 a. de C. Fue ese un período en el que se constata un rápido cambio en la Europa de clima templado y se asiste al crecimiento del comercio y la expansión de la producción metalúrgica. El conocimiento arqueológico de tal transición ha procedido principalmente de la excavación de necrópolis, cuyo número se cuenta por miles, y del casual (pero no por ello infrecuente) hallazgo de tesoros enterrados. A pesar de esto, sólo se han investigado sistemáticamente unos cuantos poblados europeos pertenecientes a tal período; de ahí que se conozca poco los procesos económicos que impulsaron la formación de las primeras ciudades de Europa Central. Describiré aquí los resultados obtenidos durante cuatro campañas de excavación llevadas a cabo en un asentamiento agrícola de esta etapa en la Baja Baviera, y lo relacionaré con la aparición, aquí y en otras partes de Europa, de poblados mayores: los precursores de las ciudades comerciales de la Edad Media.

El yacimiento, Hascherkeller, descansa sobre una terraza de arena y grava que constituye la orilla septentrional de un estrecho valle fluvial de los alrededores de Landshut, la principal ciudad de la Baja Baviera. La terraza se eleva unos 15 metros sobre el nivel del río (el Isar, afluente del Danubio); sus sedimentos de detritus glaciares se encuentran cubiertos por una gruesa capa de loess, un tipo de suelo sedimentario de color amarillo pálido atribuido a la deposición eólica de finales de la glaciación. A su vez, el loess aparece cubierto por una rica capa de humus, resultante del desarrollo milenario del suelo; los 40 centímetros superiores de esta capa húmica se han visto alterados por las modernas técnicas de arado de la tierra, propiciando la destrucción del suelo original en el que se produjeron los asentamientos prehistóricos; sólo quedan formas resultantes de las labores de excavación llevadas a cabo por los propios habitantes del asentamiento, tales como los hoyos de variada tipología y las zanjas de delimitación, observándose una intrusión del humus de relleno, de color oscuro, en los niveles loéssicos no afectados por la alteración.

A principios de 1978 abordé con mis colegas la tarea de dividir el área a excavar en cuadrículas de 5 por 10 metros. En dos excavaciones consecutivas se extrajo con palas la capa húmica alterada; en cada una de ellas se profundizó 20 centímetros. El material arqueológico que apareció entre el humus se reunió y catalogó de acuerdo con la cuadrícula y el nivel de origen. Una vez alcanzado el fondo del humus alterado, se prosiguió el trabajo con azadas y paletas hasta que apareció la capa de loess subyacente. Antes de proseguir la excavación de las formas intrusivas (hoyos y zanjas) puestas al descubierto, se cartografiaron y fotografiaron. Se bisecó el humus de relleno de cada hoyo y se examinaron por separado ambas partes. Con una malla metálica de seis milímetros de trama cribamos el humus que se extrajo de los hoyos.

A medida que el trabajo avanzaba, nos afirmábamos en la idea de que el poblado comprendía tres recintos tangenciales dispuestos en dirección este oeste y que cada uno de ellos estaba rodeado por una zanja doble. Los recintos estaban bien definidos por sus lados norte, este y oeste, pero la erosión fluvial había destruido el lado sur. Durante el primer año y las campañas que le siguieron, nuestro trabajo se centró en los recintos occidental y central.

A pesar de que el arado había destruido el primer nivel de las zanjas, la estructura subyacente tenía unos tres metros de ancho. La excavación demostró que las zanjas se habían cavado en forma de *V*, hasta un metro y medio de profundidad. El relleno, de color marrón oscuro, contenía pequeños fragmentos cerámicos y de huesos de animales. En el fondo de la zanja interior del recinto oriental había una hilera de 19 postes; evidentemente allí había existido alguna vez una empalizada. La distancia media que separaba los centros de los postes era de 13,7 centímetros. Si ello representaba el diámetro medio de los postes, se trataba, sin duda, de una empalizada bastante consistente.

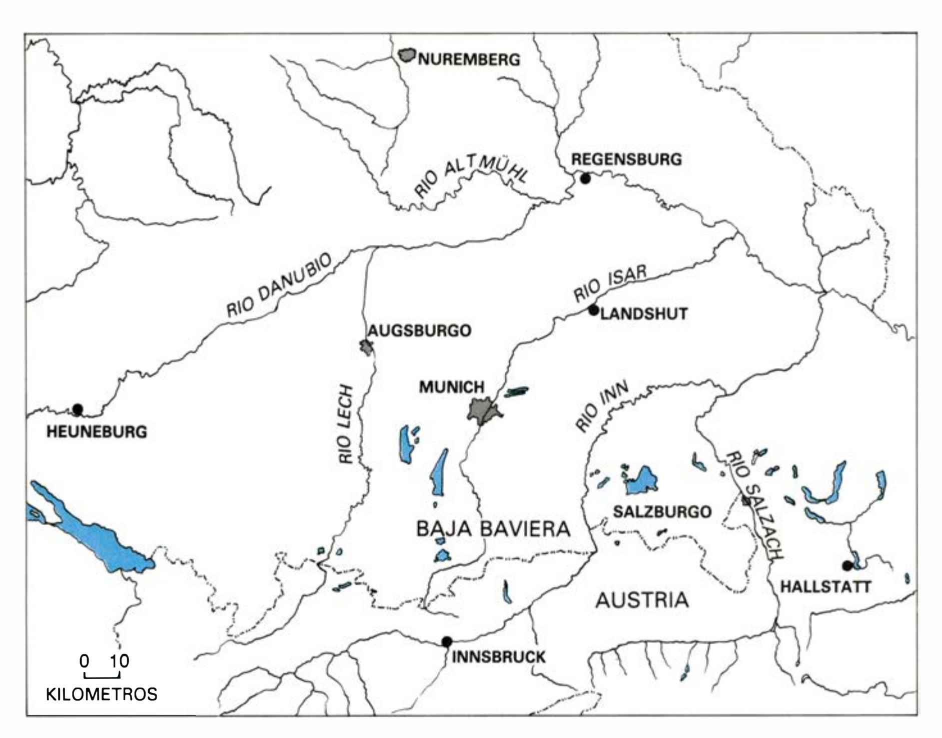
La experiencia demuestra que ese tipo de postes, clavados en la tierra, se pudren en pocas décadas; a pesar de que no se han encontrado más hoyos de postes, en muchos sectores de las zanjas se constatan indicios de reexcavaciones que habrían sido necesarias para colocar nuevos postes en sustitución de los viejos que caían. Esto apoya la teoría de que todas las zanjas contaron con empalizas, cuyo objeto



1. EXCAVACION RECTANGULAR en uno de los tres asentamientos agrícolas de Hascherkeller, yacimiento de la Edad del Hierro sito en la Baja Baviera. La distinta coloración del suelo justo delante del jalón graduado de dos metros revela parte del sistema de la zanja que separaba este asentamiento de su vecino occidental. Los agricultores erigieron empalizadas en las zanjas con el objeto de guardar el ganado y defenderse de las fieras por la noche.



2. PERFIL DE UNA ZANJA cavada en Hascherkeller hace tres milenios; está señalado por una intrusión de humus negro en forma de *V* en el loess amarillento que cubre la grava de la terraza fluvial. Esta zanja es una de las dos que, dispuestas concéntricamente, rodeaban al poblado más occidental. En las zanjas se encontraron fragmentos de huesos de animales, trozos de barro cocido que enlucía las paredes de las casas y casquetes de cerámica.



3. SITUACION DEL POBLADO. Se encuentra por encima del río Isar, en las afueras de Landshut. Cerca están Hallstatt, yacimiento industrial de la Edad del Hierro, y Heuneburg, una ciudad comercial.

sería, presumiblemente, no tanto defender el recinto de posibles atacantes cuanto guardar los animales domésticos dentro del poblado y defenderse de las fieras por las noches.

La mayor parte de la información que poseemos sobre Hascherkeller se apoya en el contenido de los hoyos que hemos encontrado dentro de los recintos. De los 21 pozos mayores que hemos excavado, 11 podrían ser asignados, en virtud de su contenido, a los tres poblados; los 10 restantes pertenecerían a poblados más pequeños de la Primera Edad del Bronce o a ocupaciones romanas posteriores. Como las zanjas, los hoyos y su humus de relleno de color oscuro resaltaban nítidamente sobre el loess subyacente. El propio relleno de los pozos era incluso más oscuro que el de las zanjas, lo que indicaba que aquéllos contenían más materia orgánica.

Atendiendo a la forma, tamaño y contenido de los 11 pozos, se pueden establecer cinco categorías funcionales. Había tres alargados, estrechos y con forma de taza; se orientaban hacia los puntos cardinales del recinto: dos en dirección norte-sur y el tercero de este a oeste. Tales orientaciones han resultado ser predominantes en las construcciones europeas pertenecientes al mismo período. La coincidencia sugiere que los hoyos servían de "bodegas": lugares destinados al almacenamiento de las grandes vasijas que protegían los víveres contra la humedad, las fluctuaciones de la temperatura y la depredación animal. Los casquetes cerámicos encontrados en los tres hoyos apoyan esta conclusión; muchos de estos casquetes eran fragmentos de vasijas de paredes gruesas.

La boca de los tres pozos mayores era circular; las paredes, casi verticales. En otros yacimientos del mismo período se han encontrado pozos similares a éstos con grano carbonizado; es más, los hoyos se habían revestido con cestería o arcilla, señal de que servían de almacén de grano. Ninguno de los tres pozos de Hascherkeller aportó pista alguna que nos revelara su función. Apenas contenían casquetes cerámicos u otro tipo de restos; quizá cumplieran una función similar.

Otros dos pozos parecen haber estado asociados a la actividad metalúrgica. En uno desenterramos un martillo de piedra y un molde de piedra arenisca para la fundición de anillos. Cerca se encontraron pequeños recortes de bronce. Otro hoyo adyacente presentaba guijarros de color rojizo y suelo aparentemente descolorido por la acción de un intenso calor. Cabe conjeturar que el fuego situado en el segundo pozo fundía el bronce para moldearlo. El pozo donde se recuperó el molde tenía también una pesa de telar, hecha de barro, y cinco torteras de arcilla. Todo ello sugiere que la zona estaba destinada tanto a la hilatura como a la fundición.

Un pozo grande, situado unos metros al norte del recinto central, contenía mucho carbón vegetal y en su fondo había restos de una estructura arcillosa. Más de la mitad de los casquetes encontrados en el yacimiento procedían de ese hoyo. Abundaban los desechos de alfarería, es decir, cerámicas que se habían reventado o deformado al exponerlas al fuego. La conclusión que parece más inmediata es que la estructura arcillosa de la que hablamos constituyese la cámara de fuego de un horno de alfarero.

Los dos últimos pozos, poco profundos y con sus lados en suave pendiente, contenían pequeños fragmentos cerámicos y huesos de animales. Aparentemente, éstos no servían al mismo propósito que los otros pozos, sino que de ellos los habitantes de Hascherkeller extraían el loess como materia prima para la fabricación de cerámica o para revocar las paredes de sus casas. Luego, esos huecos alojarían los desperdicios del poblado. En este sentido, los 11 pozos contenían un total de 198 kilogramos de barro cocido, puesto al fuego al quemarse una estructura o pegarse el barro a la pared adyacente a alguna fuente calorífica, una chimenea, por ejemplo. El resto del barro se habría secado entre el propio cieno después de que las construcciones del poblado cayeran en desuso.

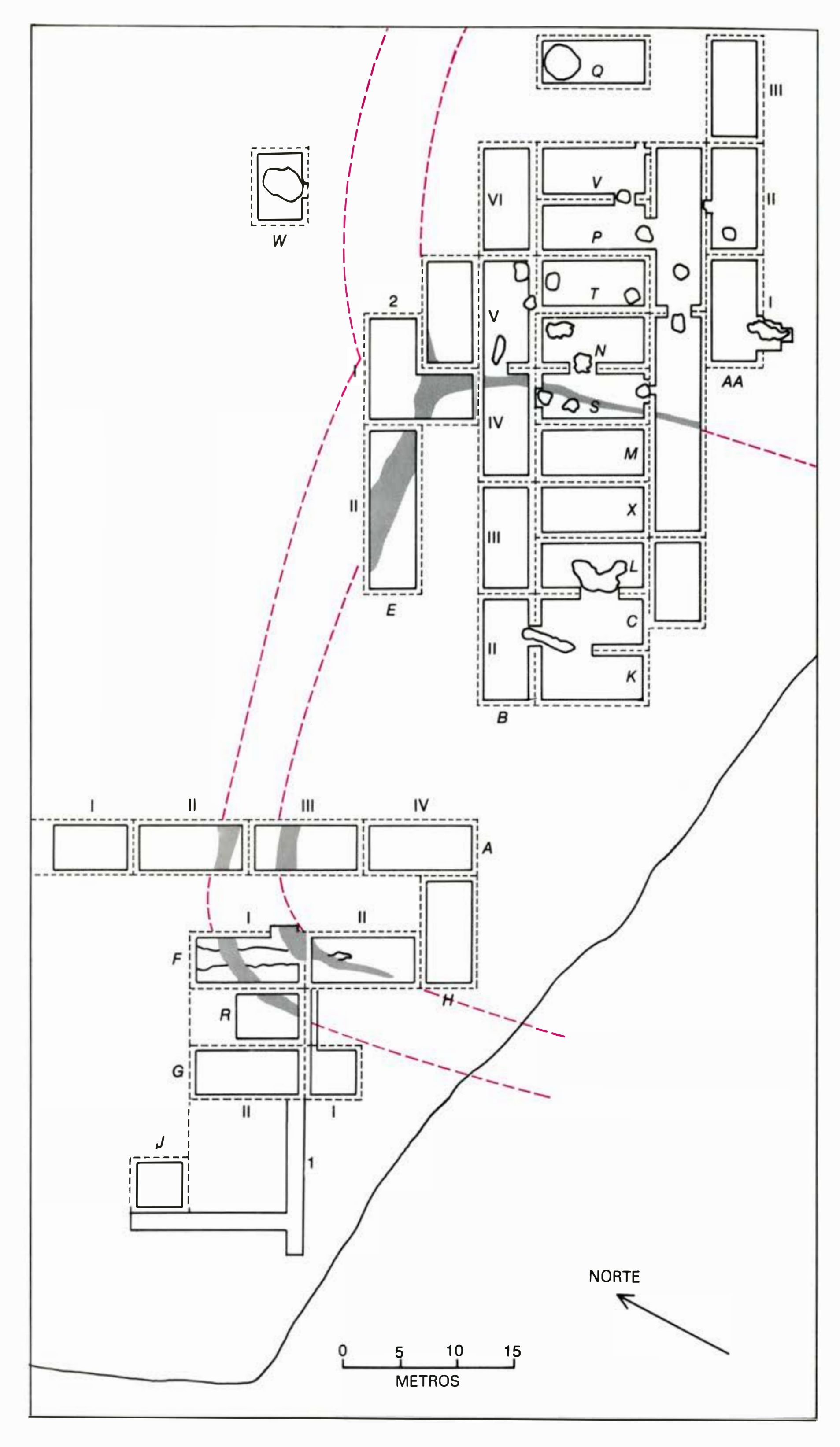
La prueba más elocuente de la ocupación humana en Hascherkeller es la cerámica deteriorada. El total de cascotes recogidos es de 14.853. Aunque la gran mayoría de ellos se encontraron en los pozos, 3828 proceden de las zanjas, de las capas húmicas que cubren el loess y de otras zonas. Todos son, de hecho, artículos típicos de la vajilla de los agricultores. Muchas tumbas de Europa central pertenecientes a esta etapa ofrecen vajillas finamente decoradas; pero en nuestro caso, los cascotes decorados no alcanzaban el 4 por ciento del total recuperado.

Clasificamos los cascotes en razón de su grosor. Podían agruparse en tres categorías principales. La primera la componían las vasijas de paredes relativamente delgadas, cuyos cascotes tenían menos de 4,5 milímetros de espesor. Formaban la segunda aquéllos cuyo grosor iba de los 4,5 a los nueve milímetros; en la tercera categoría se englobaban los cascotes que sobrepasaban los nueve milímetros. Los restos de tazas pequeñas, copas y cuencos entra-

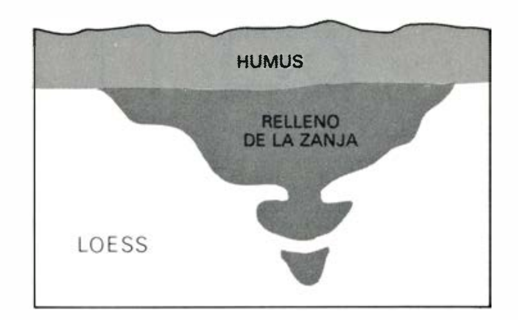
ron dentro de la primera categoría; eran también los menos numerosos y los que más a menudo presentaban decoración. Los cascotes de los cuencos más grandes y altos, jarras de boca ancha con una superficie vasta, componían la segunda categoría. Los representantes de la tercera eran principalmente jarras de superficie áspera, que probablemente sirviesen para la conservación de alimentos. En otros yacimientos del mismo período se han encontrado, enterradas en pozos de almacenamiento y "bodegas", muchas vasijas intactas que aún contenían grano.

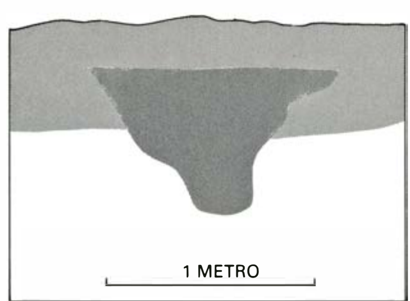
Al margen de las grandes cantidades de barro cocido que nos aportaron una valiosa información sobre la localización de estructuras dentro del poblado, los vestigios más numerosos fueron los fragmentos de huesos de animales; había un total de 1435. Del análisis efectuado por Brenda Benefit, que prepara su doctorado por la Universidad de Nueva York, se deduce que 253 de ellos pueden identificarse como pertenecientes a partes específicas del esqueleto de animales de especies conocidas. Los huesos identificados corresponden predominantemente a animales domésticos (87 por ciento); el resto son huesos de animales salvajes. Entre los huesos de animales domésticos destacan los de cerdo (37 por ciento), le siguen los de ganado vacuno (24 por ciento) y los de ovejas y cabras que, juntos, suman casi el total de los restantes (33 por ciento). También se encontró un pequeño número de fragmentos de huesos de caballo y perro. Entre los animales salvajes destaca el ciervo común (*Cervus elaphus*), pero también hay huesos de liebre, erizos y de una especie de ave, la codorniz. Menudean los huesos de peces, pero ninguno permite efectuar la identificación de la especie.

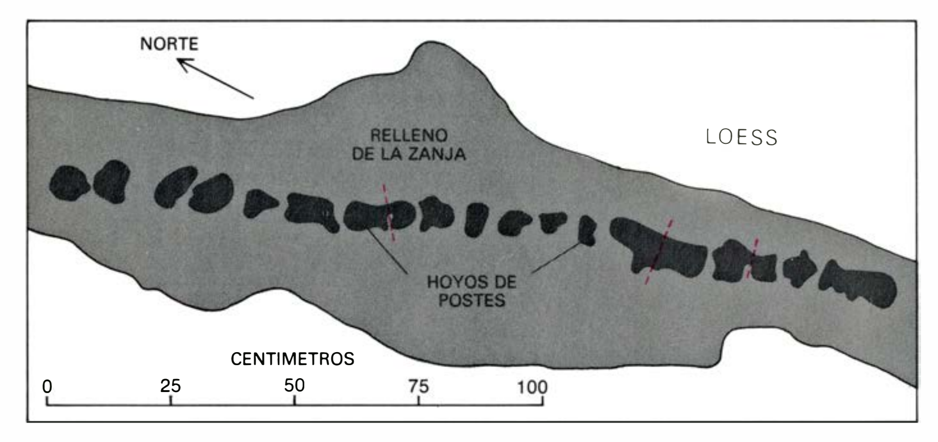
Benefit ha analizado también las piezas dentarias descubiertas en el yacimiento. El desgaste de las piezas de cerdo recuperadas ofrece un modelo doble, indicio de que los animales se sacrificaban en dos momentos diferentes del ciclo vital. Muchos se consumían lechones, es decir, al poco de su nacimiento. El resto se sacrificaba cuando tenían unos dos años. Es esta una práctica de consumo del cerdo que pervive en Europa hasta nuestros días y que permite obtener el máximo rendimiento en carne con respecto a la cantidad de alimento que debe suministrársele al animal cuando, en los meses inverna-



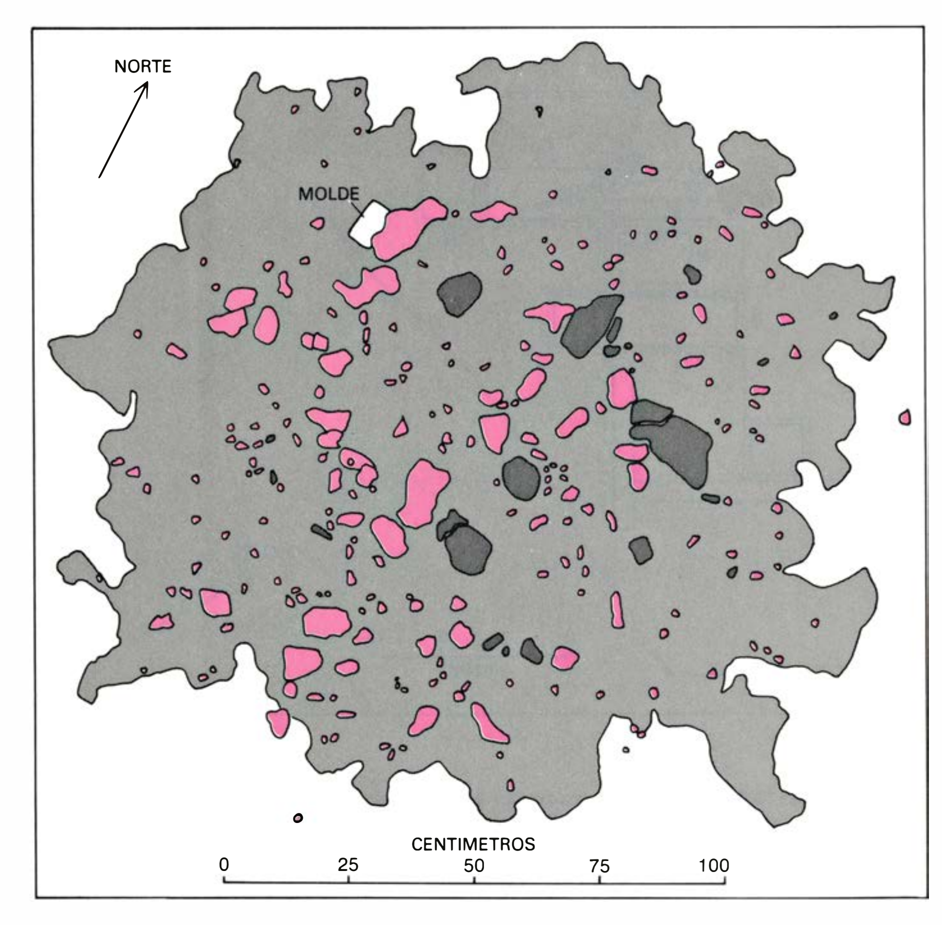
4. DOS DE LAS TRES GRANJAS, la occidental y la central, aparecen en este plano. Las partes excavadas de las zanjas dobles que rodeaban los asentamientos aparecen en gris; su extensión, indicada por una inspección de magnetómetro, está trazada con líneas discontinuas de color. De los 21 grandes pozos (*trazos continuos*) descubiertos por los excavadores, 11 pertenecían a la Edad del Hierro. Más de la mitad de los fragmentos cerámicos encontrados en el yacimiento proceden de un solo pozo (*W*) situado al otro lado de las zanjas del recinto central, que también contenía restos de un horno cerámico. Dos pozos "bodega" destacan entre los otros (*C, AA I*) por la cantidad de fragmentos de emplastes de barro que contenían, respectivamente 75 y 44 kilogramos. Del pozo del oeste que aparece en el rectángulo *N* (*véase la figura 6*) se recuperó un martillo de cuarcita, un molde de piedra arenisca para la fundición de anillos de bronce, una pesa de telar de arcilla, cinco husos y 20 kilogramos de fragmentos de emplastes de barro.







5. ZANJAS DE SECCION TRANSVERSAL EN FORMA DE *V* (*arriba, izquierda y derecha*); destacan del subsuelo de loess por el color oscuro del humus de relleno. El plano de un sector de una zanja (*abajo*) muestra la existencia de 19 agujeros de postes en un segmento de 1,6 metros. Los trazos discontinuos de color señalan hoyos dobles. Esta sección es una parte de la zanja interior del recinto occidental.



6. "TALLER" del rectángulo *N*. Contenía, además del equipo de fundición e hilatura, numerosos casquetes cerámicos (*negro*) y fragmentos de emplastes de barro (*color*) en suficiente cantidad para suponer que el "taller" de los artesanos del metal y textiles tenía también las paredes enlucidas. El molde de piedra arenisca que se utilizaba para la fundición aparece a la izquierda de la mitad superior de este plano.

les, éste no puede procurárselo por sí mismo.

Las ovejas, cabras y ganado vacuno disfrutaban de vidas más largas y se veían menos afectados por la carencia del forraje en invierno. Probablemente las ovejas se criaran para la obtención de lana y, las cabras y vacas, por la leche y sus derivados. El ganado vacuno quizá se valorase también como animal de tiro y, tras su sacrificio, se aprovecharan sus pieles.

Además de esta importante reserva de carne y productos lácteos, los restos de diversas plantas encontradas en Hascherkeller demuestran que los agricultores contaban con el cultivo de cereales y ciertas plantas silvestres para completar su dieta alimentaria. El análisis de los restos de plantas realizado por Caroline Quillian Stubbs, doctoranda por la Universidad de Harvard, señala que el principal cereal cultivado era el mijo, además del trigo y la cebada, y que también conocían las lentejas. Los habitantes de Hascherkeller recolectaban avellanas y frutos de diversas plantas silvestres, consideradas hoy como yerbajos no adecuados a la alimentación humana, pero que entonces desempeñaron un papel importante en la dieta de los primeros europeos. En este sentido, se recolectaba el galio (*Galium*), el quenodio (*Chenopodium*) y la canela (*Rumex*).

A partir de estas modestas pistas, ¿qué panorama de la vida de la Primera Edad del Hierro se nos aparece? El tamaño aproximado de los tres recintos rodeados por zanjas es de unos 3000 metros cuadrados cada uno, lo que se corresponde con el tamaño de otros muchos recintos de la última etapa de la prehistoria europea y de los primeros tiempos históricos. Es probable que cada "granja" estuviera habitada por una familia compuesta por cinco o 10 miembros y que cada unidad comprendiera una vivienda y un granero o cobertizo para almacenar los excedentes de producción, aparte de otras estructuras menores como los alpendes y talleres. Es evidente que los tres asentamientos se ocuparon al mismo tiempo: sus zanjas, contiguas sin entrecortarse, así lo atestiguan.

La vida cotidiana de los 15 o 30 hombres, mujeres y niños del poblado puede considerarse bajo tres perspectivas económicas interrelacionadas: subsistencia, manufactura y comercio. De la primera existen pruebas evidentes que nos hablan de una economía de autoabastecimiento: por un lado el pasto-

reo de animales, que proporcionaba carne y leche, complementado con la caza y pesca; por otro, el cultivo de cereales y legumbres, complementado por la recolección de frutos silvestres. Independientemente de que estas actividades agrícolas ofreciesen un excedente de carne o grano, es probable que otras actividades, como la fabricación de queso o la elaboración de vestimentas de cuero, proporcionasen al poblado artículos de fácil conservación que excedieran las propias necesidades de subsistencia.

¿Qué otros artículos de consumo, además del queso, cuero y posiblemente carne, pudieron arrojar excedentes? No parece que la cerámica lo fuera; habida cuenta de que los cuencos, tazas y jarras no se hacían en un torno, sino mediante una ardua tarea de moldeado a mano que difícilmente pudo ofrecer al poblado un excedente cerámico. Nuestro hallazgo de pesas de telar y torteras, junto con el indicio de que las ovejas se criaban durante bastantes años antes de su sacrificio, sugieren que la producción de vestimentas de lana estaba por encima de las necesidades locales.

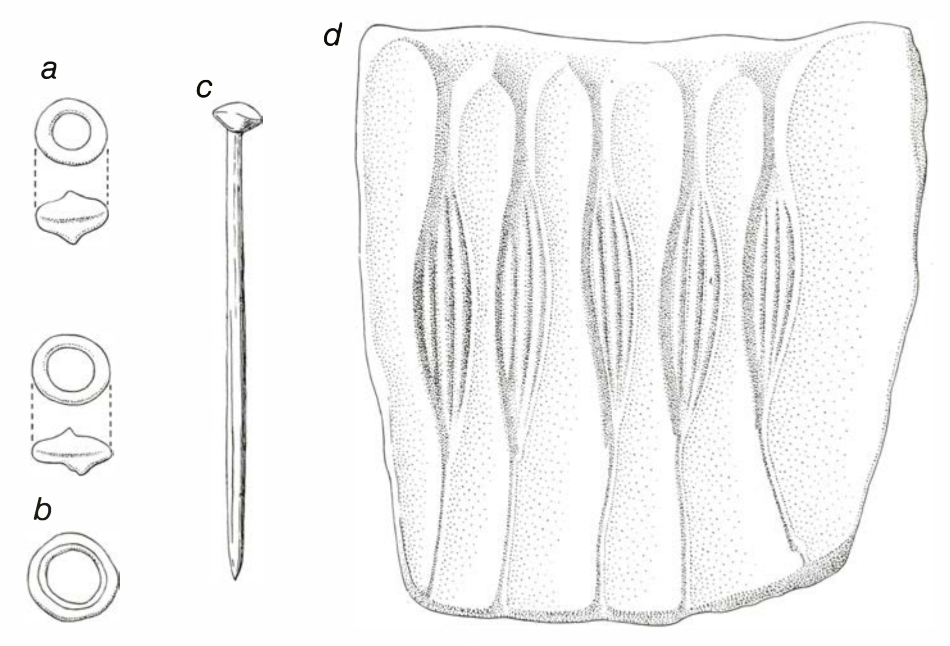
Llegamos así a la tercera actividad económica: el comercio. Prueba de que los habitantes del poblado importaban artículos exóticos la tenemos en los objetos de bronce (anillos, por ejemplo), que se hacían de recortes de bronce fundido. Sin duda, los recortes de bronce eran de importación, y varios alfileres encontrados en el poblado tampoco parecen autóctonos. Por otro lado, también encontramos cinco fragmentos de objetos de hierro, demasiado pequeños y corroídos para identificarlos, que, aunque testifican que el poblado es de la Edad del Hierro, no implican necesariamente que los agricultores importasen objetos de dicho metal. El mineral de hierro abunda casi por doquier; no así las materias primas para la aleación que conocemos por bronce, en concreto el estaño. La prueba de la fundición local del hierro nos la aporta la escoria encontrada en el yacimiento, lo que abona la conclusión de que las gentes de Hascherkeller estaban familiarizadas con la forja. No existe, por contra, el menor indicio de fundición local de bronce.

Entre los objetos importados cabe incluir también las cuentas de pasta de vidrio y el grafito, éste utilizado para decorar vasijas de manufactura local. Si el grafito se aplica a las vasijas antes de someterlas a la acción del fuego, adquieren una superficie negra y brillante. Tal tipo de ornamento, tanto si reviste la superficie del vaso en su totalidad como si lo hace formando series de bandas, era algo muy extendido por Europa central a principios del año 1000 a. de C. Los principales centros abastecedores de grafito se encontraban, por lo menos, a 100 kilómetros de Hascherkeller, al este de Passau, sobre el Danubio, y al norte, en Bohemia. Respecto a las cuentas de vidrio que recuperamos (cuatro de ellas de color azul verdoso), dos procedían del hoyo que contenía el horno cerámico y otras dos de otro. Desconocemos su lugar de fabricación. Ni en nuestro yacimiento ni en otros poblados de Europa central del mismo período se ha encontrado la menor señal en este sentido. Sin embargo, a juzgar por la relativamente extensa distribución de estas cuentas, parece lógico pensar que existiese más de un centro productor y de exportación a las comunidades agrícolas de la zona.

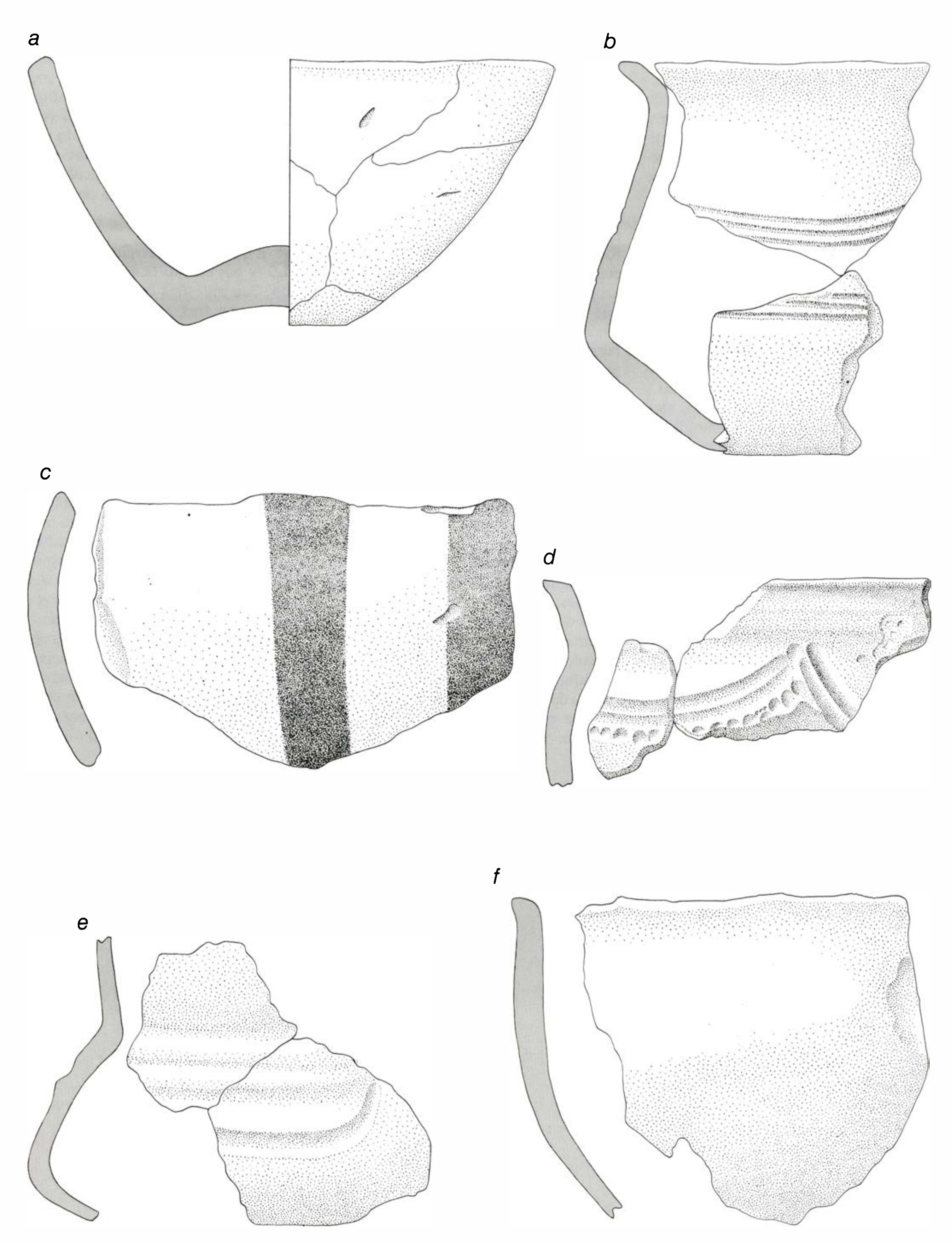
Lo que convierte a este modelo de importación de bronce, grafito y cuentas de vidrio en algo particularmente significativo en relación con el posterior desarrollo de las ciudades de la Edad del Hierro de la Europa situada al norte de los Alpes es que los tres son artículos de lujo. En Hascherkeller, la manufactura de hoces de metal y otros aperos de labranza, si bien puede considerarse quizá demasiado preciosista, tenía sin duda un fin utilitario. Sin embargo, ningún agricultor necesitaba, para poseer más cerdos o plantar más mijo, anillos y alfileres de bronce, ni tener la cerámica decorada con grafito, ni cuentas de vidrio de color azul verdoso. Es evidente que el poblado comerciaba con el mundo exterior para atender a la demanda de artículos de lujo, y no para satisfacer sus necesidades primarias.

¿En qué fecha ocurrían estas cosas y por cuánto tiempo se prolongaron? Seis muestras de carbón vegetal recogidas en tres pozos, sometidas a datación por medio de carbono 14, indicaron que la ocupación del poblado tuvo lugar en un período de 200 años comprendido entre el 1000 y el 800 a. de C. A su vez, los objetos de bronce, la cerámica y las cuentas de vidrio importadas coinciden con la fase arqueológica que para Europa central se conoce con el nombre de Hallstatt *B*. Cuando se realiza una datación comparativa del material correspondiente a esta fase con el material procedente de las culturas del Mediterráneo del que se dispone de registro histórico, la fase *B* del Hallstatt se sitúa en los tres siglos que van del 1000 al 700 a. de C. Así pues, tanto la cronología absoluta como la relativa están en estrecha concordancia.

Respecto a la cuestión de cuánto tiempo perduró el poblado, las pruebas son menos concluyentes. Las zanjas se renovaron, y ya se ha indicado que la



7. ARTICULOS DE LUJO hallados en Hascherkeller: cuentas de vidrio (*a*), un anillo de bronce (*b*) y un alfiler, también de bronce (*c*). Aunque se halló el molde para la fundición de las láminas de bronce (*d*), no se halló ningún anillo que procediese de él. El alfiler y el anillo, así como las cuentas de vidrio y el grafito utilizado para la decoración de la cerámica, probablemente se trocaran por excedentes agrícolas.



8. CASCOTES DE SEIS VASIJAS que forman, parcial o completamente, el perfil original de los vasos. Los cascotes decorados son más finos que los que no lo están (*a*); una vasija (*c*) se decoró con bandas negras conseguidas por la aplicación del grafito antes de su horneado. Los centros de abastecimiento de grafito más cercanos se hallaban a 100 kilómetros, por lo que los alfareros del lugar posiblemente se procuraran este artículo a través del comercio.

renovación era necesaria para la sustitución de los postes inservibles. Partiendo del hecho de que las empalizadas podrían aguantar unas pocas décadas, el poblado no pudo durar menos. El barro del enlucido de las paredes conservado por cocción nos ofrece también otra pista sobre el particular. Muchos trozos de barro indican que las estructuras habían sido enlucidas y vueltas a pintar dos o tres veces, lo que revela que la ocupación duró dos o tres generaciones. ¿Pudo haberse prolongado más? Tanto los datos aportados por el carbono 14 como los análisis de la cerámica indican que la ocupación no superó el límite de los dos siglos. La cerámica es de un estilo totalmente homogéneo y resulta difícil creer que el poblado se mantuviera al margen de las corrientes estilísticas europeas por un período superior a los 200 años.

La imagen que después de cuatro campañas de excavación poseemos de Hascherkeller es la de una comunidad económicamente autosuficiente, vinculada, al mismo tiempo, al ámbito centroeuropeo por medio de la apetencia de pequeños lujos. Para satisfacer tal apetencia sólo había un camino: el comercio. Y éste era bidireccional; los agricultores ofrecían sus productos agropecuarios: vestidos de lana o hilo, queso, posiblemente manteca, carne sazonada y cuero (o artículos de cuero curtido).

Por ese tiempo se estaba gestando en Europa central una intensificación general de la producción agrícola. Al margen de que el poblado de Hascherkeller perviviera durante dos siglos, otros asentamientos agrícolas duraban tanto o más tiempo. El cultivo de rotación, el abonado y barbecho de la tierra conservaban la fertilidad del suelo; por otro lado, se fabricaban ya muchas herramientas metálicas: hoces, hachas, sierras, cinceles y martillos, útiles todos ellos que incrementaban la eficacia de la acción del hombre sobre su entorno.

Imaginemos por un momento que la expansión de la metalurgia contribuyera a la formación de asociaciones de forjadores que no cultivaban la tierra. Tal y como revelan las tumbas del período, no sólo se forjaban y fundían herramientas, sino que el difunto se hacía acompañar también con armas decoradas y artículos de valor: espadas, cascos, grandes vasijas de bronce y ornamentos de oro. ¿Cómo se procuraban los recursos alimentarios? Probablemente cambiando sus manufacturas por los excedentes agrícolas de los asentamientos agropecuarios, que cada vez incrementaban más su producción. Podemos incluso imaginar la existencia de viajantes intermediarios insertos en esta cadena de intercambios.

El registro arqueológico ofrece un ejemplo real de una asociación similar que apareció en lo que hoy es Austria, a no más de 160 kilómetros de Hascherkeller. Una comunidad, localizada en las minas de sal de Hallstatt, entre el 800 y el 400 a. de C., formada por unos 200 individuos, dedicó sus esfuerzos de modo exclusivo a la extracción y comercio de la sal. El conjunto de artículos, verdaderamente excepcionales, que acompañaba a los difuntos de Hallstatt, constituyen una prueba elocuente del éxito alcanzado en esta empresa. Los primeros mineros de Hallstatt comenzaron la explotación hacia el 1000 a. de C., probablemente extrayendo la sal para su propio consumo, al igual que harían los primeros fundidores de bronce, tejedores, pastores o fabricantes de queso de Hascherkeller en su pequeño poblado de los siglos posteriores.

En cualquier caso, la aparición de la ciudad minera de Hallstatt no fue un episodio aislado. Al pie de los Alpes, en la actual Eslovenia, se levantaron ciudades habitadas por cientos de individuos dedicados a la fundición y forja del hierro. Cerca de Hascherkeller, al suroeste de Alemania, en un yacimiento bien estudiado, se estableció un centro comercial: Heuneburg; un recinto de sólidas construcciones de madera no muy distinto del que sirvió para las primeras ciudades comerciales del medioevo de 14 siglos después. Los artesanos de Heuneburg, más que desempeñar una actividad concreta, se especializaron en diversos tipos de manufacturas, y lo mismo ocurrió en otras ciudades comerciales de la Europa central. Sin embargo, la principal función de las ciudades fue la comercial. Ninguna de esas poblaciones habría florecido sin el apoyo de miles de pequeños poblados agropecuarios como el de Hascherkeller, capacitados y dispuestos a producir siempre más excedentes para canjearlos por artículos de las ciudades. En *La Tempestad*, de Shakespeare, Antonio dice: "¿Qué es el pasado?, un prólogo." En este mismo sentido, podemos ver en Hascherkeller que la Primera Edad del Hierro fue un prólogo al urbanismo medieval y renacentista que, en última instancia, formaron nuestro mundo moderno.

Ciencia y sociedad

Nobel de medicina 1983

El premio Nobel de medicina 1983 ha sido concedido a la científica norteamericana Bárbara McClintock, de 81 años de edad, por su trabajo sobre los elementos genéticos móviles. Iniciado en el ámbito de la investigación más pura, menos aplicada, hace ya más de cuarenta años, sólo ahora ha recibido el reconocimiento oficial de la comunidad científica. McClintock demostró la existencia de estos elementos en el genoma del maíz y explicó con ellos la extraordinaria variegación en el color de sus semillas. Recientemente se ha descubierto la presencia de elementos similares en organismos muy diversos.

Acuñó ella la expresión "elemento controlador" para designar al capaz de desplazarse de un lugar a otro en los cromosomas y regular la actividad de los genes "normales" (estructurales). Un elemento controlador puede inhibir la actividad de un gen al integrarse en el interior del propio gen (interrumpiendo, por tanto, su mensaje) o en un lugar cercano al mismo. De vez en cuando, el elemento controlador se escinde. Entonces, la actividad del gen se recupera en mayor o menor grado, mientras que el elemento se reinsertará quizás en otro punto del genoma para afectar la actividad de un gen distinto. Deleciones, duplicaciones, inversiones y translocaciones de material genético pueden ocurrir en los sitios donde se albergan elementos controladores. La rotura de cromosomas constituye otra consecuencia común de su presencia.

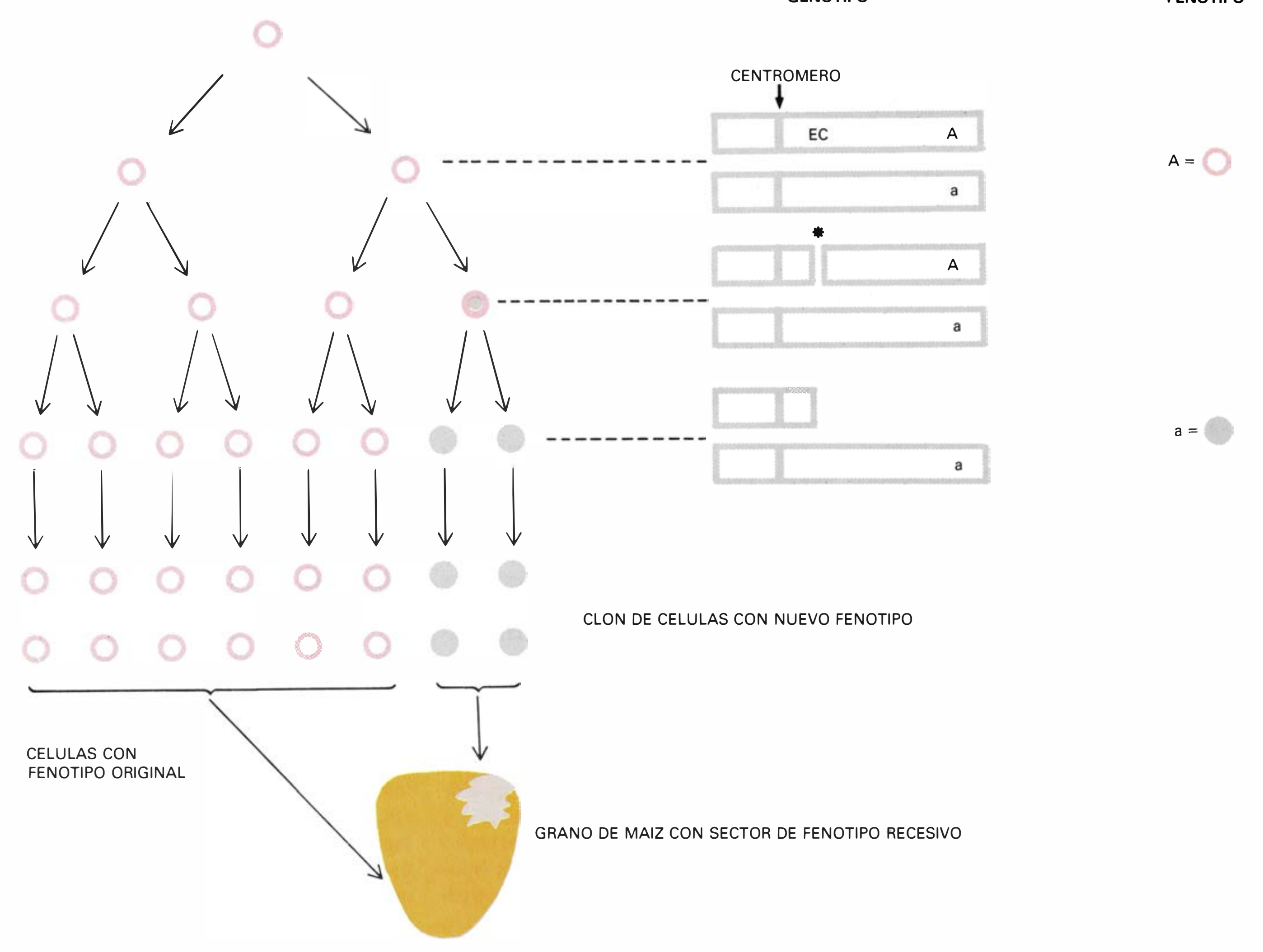
En el maíz se distinguen dos clases de elementos: entidades sin productos genéticos conocidos, denominadas elementos receptores, y unidades más complejas, los elementos reguladores. Estos últimos pueden sufrir integración, escisión y transposición, procesos para los cuales se postula la existencia de enzimas específicos, capaces de catalizarlos. Tales enzimas parecen ser suministrados únicamente por los elementos reguladores, ya que sólo éstos pueden moverse autónomamente y controlar, además, el movimiento de los elementos receptores. En otras palabras, los elementos receptores son incapaces de desplazarse sin la presencia de un elemento regulador en alguna parte del genoma.

La combinación de una unidad reguladora activa en "trans" y de un elemento receptor no autónomo constituye un sistema de dos elementos. Se han encontrado varios de ellos en el maíz, siendo los sistemas Ac-Ds, Spm y Dt los mejor caracterizados. La relación receptor-regulador es altamente específica, lo que significa que un elemento receptor sólo puede ser activado en "trans" por su correspondiente unidad reguladora. En un sistema autónomo, o sistema de un elemento, regulador y receptor son partes integrantes de la misma entidad, residiendo ambos en el mismo *locus* bajo control. Parece muy probable que los elementos no autónomos deriven de sistemas autónomos tras la pérdida de las funciones capaces de actuar en "trans", necesarias para el desplazamiento o transposición. Sabido es que reguladores y receptores cambian frecuentemente sus propiedades. Estos cambios, aunque tranmisibles como mutaciones a través de muchas divisiones celulares (hereditarios, por tanto) se les denomina a veces cambios de "estado" o de "fase", en razón de su alta frecuencia, su fuerte tendencia a revertir y (en algunos casos) la facilidad con que se pueden predecir en ciertos estados del desarrollo de la planta.

Dos rasgos del maíz han facilitado el estudio de los sucesos de transposición. Por un lado, los elementos controladores suelen insertarse cerca de genes que tienen efectos visibles, aunque no letales, sobre el fenotipo: la pigmentación de las semillas, por ejemplo. Por otro, el desarrollo clonal del maíz permite visualizar la aparición y cronología del suceso de transposición. La figura adjunta nos ilustra el proceso. No importa cuál sea la naturaleza del suceso en cuestión: inserción, escisión, rotura cromosómica o cualquier otro. Pero sí se requiere que ocurra en un heterozigoto, de manera que la expresión de uno de los dos alelos quede alterada. Así, los descendientes de una célula donde se ha desarrollado el suceso mostrarán un nuevo fenotipo, mientras que los descendientes de las células no afectadas continuarán con el fenotipo original. Los descendientes mitóticos de una célula constituyen un clon; permanecen en el mismo lugar produciendo un sector del tejido. Un cambio en el fenotipo durante el desarrollo somático se denomina "variegación" y se revela como un sector de nuevo fenotipo sobre un fondo de tejido con fenotipo original. El tamaño del sector depende del número de divisiones a partir del momento en el que se produce la transposición.

Bárbara McClintock desarrolló fundamentalmente su trabajo sobre los elementos móviles del maíz en la década de los cuarenta, mientras trabajaba en el Carnegie Institute, con sede en Cold Spring Harbor. Tras varios años de intensa investigación reunió sus principales resultados y conclusiones y los presentó en el simposio que, en 1951, tuvo lugar en Cold Spring Harbor. Entre sus colegas encontró una casi total incomprensión. Pocos entendieron lo que estaba diciendo y eran menos los preparados para aceptarlo. La teoría de la movilidad de genes contradecía las doctrinas de la época, que consideraban, a los genes, estructuras fijas y estáticas dentro del cromosoma, a manera de perlas en un collar, analogía utilizada en más de una ocasión. McClintock no se limitaba a proponer la existencia de elementos genéticos cuyo desplazamiento controlaba la expresión de otros genes; postulaba que tales reorganizaciones del material genético debían desempeñar un importante papel durante el desarrollo del organismo. Ideas que laboraban contra una ortodoxia, a la que, por otra parte, ella misma había prestado antes firme apoyo. En efecto, en 1931 había demostrado que la recombinación genética durante la meiosis iba acompañada de una transferencia física de material entre cromosomas homólogos. Es decir, que al sobrecruzamiento (o "crossing-over") genético le acompañaba un sobrecruzamiento citológico. Esta conclusión, base fundacional de la citogenética, se convirtió en prueba determinante de que los genes eran parte de los cromosomas.

En 1953 volvía a insistir en su teoría de la movilidad de genes a través de un artículo publicado en *Genetics*, artículo que tampoco mereció la debida atención de sus colegas. En el simposio de 1956 de Cold Spring Harbor, dio una conferencia bajo el título "Controlling Elements and the Gene". El resultado fue el mismo que en anteriores ocasiones. Cuando McClintock leyó el artículo de Jacob y Monod sobre el operón,



1. Desarrollo clonal del maíz. La rotura (*) producida por un elemento controlador (EC) conduce, tras la mitosis, a la pérdida de un fragmento cromosómico. Si el fragmento es portador de un alelo dominante (A), su pérdida acarreará un cambio en el fenotipo. Las células descendientes de aquella en la que se produjo la rotura cromosómica muestra un fenotipo a *diferente al resto, que conserva el fenotipo original* A.

publicado en francés a finales de 1960, llegó a la conclusión de que el sistema de control genético que éstos habían puesto de manifiesto en *Escherichia coli* era similar al que ella había observado en el maíz. Analizando los paralelismos entre los sistemas de control en maíz y bacterias escribió un artículo para *The American Naturalist*, del que envió una copia a Monod en la primavera de 1961. Tampoco Jacob y Monod habían reparado hasta entonces en los trabajos de McClintock, pues no los citaban en su revisión sobre el modelo del operón que, pocas semanas más tarde, apareció en el *Journal of Molecular Biology*. Pero rectificaron y reconocieron las interesantes contribuciones de McClintock, y así lo comunicaron en el simposio que, en Cold Spring Harbor, tuvo lugar en el verano de aquel mismo año.

Hubo que esperar, pues, a la década de los sesenta para que llegara a admitirse que las ideas de McClintock sobre los elementos genéticos móviles trascendían la genética del maíz. Tuvo en ello que ver la acumulación de resultados que evidenciaban la presencia de elementos similares en numerosas especies. Primero se comprobó, en el plano molecular, la existencia de elementos móviles de bacterias y virus. Son elementos móviles los que explican que una bacteria desarrolle resistencia a un antibiótico o que un parásito como el tripanosoma, causante de la enfermedad del sueño, altere las propiedades de su membrana y escape así a las defensas inmunológicas del hombre y los animales. Elementos móviles se han encontrado también en *Sacharomyces cerevisiae* y en *Drosophila melanogaster*. Más recientemente, la existencia de estos elementos se ha puesto de manifiesto en organismos superiores, y se ha confirmado a nivel molecular, usando técnicas muy refinadas, lo que McClintock dedujo con métodos rudimentarios. Tenemos, por último, pruebas del papel crucial que ciertas reorganizaciones del material genético desempeñan en estados fisiológicos normales (producción de anticuerpos, por ejemplo) y patológicos (desarrollo de ciertos cánceres).

Al hacer público su nombramiento, la Academia Sueca ha descrito el trabajo de Bárbara McClintock como un trabajo que "ha ido por delante de su tiempo" y ha mencionado el paralelismo de la recién galardonada con Gregor Mendel, descubridor de las reglas básicas de la herencia. Ambos introdujeron conceptos que revolucionaban las ideas de la época, ignorados por sus más directos contemporáneos. Parece indudable que los partidarios de T. S. Kuhn encontrarán en el caso de

McClintock un nuevo punto de apoyo en la defensa de su teoría sobre las revoluciones científicas. Quizás un factor que adicionalmente pudo influir en la falta de receptividad ante las ideas innovadoras de B. McClintock fuese su condición de mujer en una comunidad como la científica dominada, aún más que hoy, por hombres. No hay que olvidar que a lo largo de su carrera encontró por esta causa serias dificultades. Bárbara McClintock ha sido la tercera mujer en obtener el premio Nobel de medicina desde su fundación en 1901 y la primera en conseguirlo no compartido.

A pesar de los muchos galardones recibidos en los últimos años, McClintock conserva sus hábitos de persona solitaria y humilde. Vive espartanamente cerca del laboratorio donde trabaja la mayor parte del día, todos los días de la semana. No tiene colaboradores, auxiliares o secretarias y confiesa, tras conocer su nombramiento, que "quizá parecerá injusto premiar a una persona por haber disfrutado durante tantos años haciéndole preguntas al maíz y luego contemplando sus respuestas". (María Eugenia Armengod.)

Nobel de economía 1983

La economía, aunque se considere una disciplina relativamente joven, tiene ya una andadura de más de 200 años como ciencia independiente. Hace ahora 208 años que se publicó en Gran Bretaña su partida oficial de nacimiento: *An Inquiry into the Nature and Causes of the Wealth of Nations*, producto de las reflexiones de Adam Smith, un catedrático escocés de lógica y de filosofía moral.

Naturalmente, las esperanzas de Adam Smith de que la economía de los países funcionase como un delicado mecanismo que, bajo la supervisión de una especie de "mano invisible", establecía los equilibrios necesarios para que se produjesen las cantidades adecuadas a los precios justos, fueron frustrándose con el tiempo. A mediados del pasado siglo, uno de los más influyentes escritores británicos de la época, Thomas Carlyle, atacaba con vehemencia el nuevo orden industrial, denunciaba la filosofia del "laisser faire" y calificaba de ciencia funesta o lúgubre ("dismal science") a la todavía titubeante ciencia de la denominada economía política.

Pocos años más tarde (1874) y casi coincidiendo con el primer centenario de la obra de Adam Smith, un ingeniero de minas y periodista fracasado, el economista francés Marie Ésprit Léon Walras, publicaba sus *Eléments d'économie politique pure*, en que planteaba los principios del análisis del equilibrio general. En 1983 la Academia de Ciencias de Suecia acordó conceder el premio Nobel de economía, en su quinceava edición, a otro francés, éste residente y nacionalizado en los Estados Unidos: Gerard Debreu. Sin embargo, los trabajos de Debreu, profesor en la Universidad de California en Berkeley, y los de Walras en la Universidad de Lausana, siguen una línea común de preocupación, con la natural distancia que separa el instrumental de análisis del economista con un siglo de diferencia.

Gerard Debreu ha dedicado la mayor parte de su obra a analizar el problema del equilibrio general planteado por Walras, a determinar en qué condiciones se produce un equilibrio simultáneo de precios y cantidades en los distintos mercados del sistema económico en su conjunto. Su obra clave, escrita originalmente en inglés, fue publicada en 1959 con el título *Theory of Value. An Axiomatic Analysis of Economic Equilibrium*, como monografía n.º 17 de la Fundación Cowles, en una serie que ha recogido aportaciones claves en otros dominios de la economía y de la metodología científica, tales como la econometría o la investigación operativa. Existe una traducción al castellano con el título "Teoría del valor. Un análisis axiomático del equilibrio económico". (Barcelona, Bosch, 1973).

Antes de la publicación de esta monografía, Debreu había ya apuntado en diversos artículos y en su propia tesis doctoral (leída en la Universidad de París) algunos de sus planteamientos fundamentales. Como curiosidad, puede citarse que en 1954 y 1958 publicó en la revista *Econometrica* dos artículos en colaboración con el también premio Nobel (1972) Kenneth J. Arrow: "Existence of an equilibrium for a competitive economy" y "On the stability of the competitive equilibrium".

La obra de Debreu se recuerda por sus aportaciones a la teoría económica del equilibrio general, pero sería injusto el reducirla a sólo este tema, por muy importante que sea. Además, Gerard Debreu ha contribuido a la actual teoría de la utilidad y de las preferencias, a los planteamientos en economía de la incertidumbre, y ha utilizado, con un rigor ampliamente reconocido, el instrumental matemático a su disposición, incluso con innovaciones en la teoría de la optimización y álgebra matricial.

Su obra es eminentemente teórica. Una formulación matemática rigurosa (en particular teoría de conjuntos y topología general) sirve de vehículo formal a un análisis metodológicamente cauto, que permite elaborar un cuerpo de doctrina consistente sobre el funcionamiento global del sistema económico, aunque sea, en cierta forma, a costa de unos supuestos simplificadores, a los que no siempre la realidad económica responde.

Así, no es fácil suponer que en un país concreto no existen unidades económicas con influencia individual en ciertos mercados; o que los consumidores elijan su composición de bienes consumidos en forma estrictamente racional según su estructura de preferencias; o que las empresas adoptan en todos los casos el plan de producción óptimo entre todos los tecnológicamente posibles.

La obra de Debreu no trata del poder de las multinacionales o los sindicatos, del comportamiento del consumidor inducido por la publicidad, de los problemas de I + D y la estrategia empresarial de nuevos productos. Pero ello no quita importancia a un planteamiento teórico integrador, permanentemente reivindicado por los economistas teóricos y aplicados. "En primer lugar la investigación del profesor Debreu se presenta como una construcción axiomática en la que lo fundamental no es tanto elaborar concepciones teóricas nuevas como *deducir rigurosamente, a partir de un sistema de hipótesis claramente especificadas*, proposiciones que, hasta entonces, no podían ser consideradas, al menos desde el rigor formal, otra cosa que conjeturas. En segundo lugar, la investigación del profesor Debreu intensifica, consolida y difunde un proceso de ampliación del conjunto de conceptos y métodos matemáticos utilizados en la investigación teórico-económica. En tercer lugar, la investigación del profesor Debreu ha contribuido poderosamente a la propagación de una determinada concepción (con independencia, esto es, del objeto específico de la investigación) de la investigación teórica-económica" [A. Rubio de Urquía, comunicación personal]. Todos aquellos que, en mayor o menor medida, nos dedicamos a la elaboración de modelos econométricos operativos que permitan explicar y predecir el comportamiento de una economía en su conjunto, sabemos de las dificultades provenientes de enfoques teóricos parciales sobre los que estos modelos se elaboran.

Los brillantes enfoques de la teoría keynesiana (acaba de cumplirse el centenario del nacimiento de J. M. Key-

nes) han permitido un desarrollo acusado de los modelos econométricos por el lado de la demanda. Ello supone disponer de una explicación coherente de lo que ocurre con la demanda de bienes de consumo, producidos en el país o importados, por parte de las familias; de la demanda de bienes de inversión por las empresas; del empleo generado y de las rentas que se deducen.

Sin embargo, como señalaba recientemente Lawrence Klein, Nobel de economía de 1980, "Alfred Marshall nos recordaba que ambas cuchillas de la tijera deben cortar, y ni el análisis de la oferta aislado ni el solo análisis de la demanda pueden proporcionarnos la adecuada explicación que se espera en economía" [*The economics of supply and demand*].

A este enfoque integrador, de considerables repercusiones en la práctica econométrica y que abre el desarrollo de temas sobre tecnología, "inputs" intermedios (por ejemplo, el tan debatido de la energía), productividad, precios,..., ayudan trabajos básicos como los elaborados por Debreu sobre el equilibrio general.

Desgraciadamente, siguen resultando hoy día parcialmente válidas aquellas afirmaciones que hace ya más de quince años hicieron Fox, Segupta y Thorbecke respecto al escaso reconocimiento social de la labor del economista teórico: "los científicos e ingenieros que diseñan vehículos espaciales están en una posición envidiable. Comienzan con teorías que un lego encontraría inalcanzables y aplican transformaciones matemáticas que encontraría incomprensibles. Pero al final colocan en órbita un vehículo que brilla como una estrella; aparece en el cielo de la noche cuando y donde ellos dijeron; y su éxito es visiblemente evidente no sólo para el propio lego, sino para su hijo de cinco años. El economista que crea una nueva teoría con la ayuda de matemática avanzada es respetado por sus propios colegas. Pero el lego permanece muy estirado, impasible a las prescripciones políticas derivadas de ella: 'Puede ser correcto en teoría, pero no funcionará en la práctica'. Este veredicto es rendido con tan completa autosuficiencia como para implicar que cualquier prescripción económica que es correcta en teoría es, ipso facto equivocada en la práctica". (A. Pulido.)

ERRATA

En el número 88, correspondiente al mes de enero de este año, los duendes de imprenta se ensañaron con el artículo dedicado a la historia del metro y su nueva definición.

			Dice:	*Debe decir:*
Pág. 42	Columna 3	Línea 28	la de la luz	la propagación de la luz
Pág. 42	Columna 3	Línea 42	diezmilésima	diezmillonésima
Pág. 42	Columna 3	Línea 45	Jorge Juan de Ulloa	Jorge Juan y Ulloa
Pág. 43	Cuadro 2	Línea 5	metro internacional	metro internacional. (Incertidumbre 1×10^{-6}.)
Pág. 43	Cuadro 2	Línea 7	es el vacío	en el vacío
Pág. 43	Cuadro 2	Línea 8	(Incertidumbre $1{,}10^{-8}$.)	(Incertidumbre 1×10^{-8}.)
Pág. 43	Cuadro 2	Línea 10	tidumbre $4{,}10^{-10}$.)	tidumbre 4×10^{-10}.)
Pág. 43	Cuadro 3	Línea 7	$127_{I_{d2}}$	127_{I_2}
Pág. 43	Columna 1	Línea 13	incertidumbre medida	incertidumbre de la medida
Pág. 43	Columna 2	Línea 9	a la del segundo.	a la del segundo, sin más que referirla a él.

La hélice del ADN y su lectura

El análisis por rayos X de cristales de tres tipos de moléculas de ADN bicatenario muestra que la secuencia de bases se refleja en la estructura local de la hélice

Richard E. Dickerson

La doble hélice del ADN contiene dos tipos de información genética, almacenados e interpretados de modos muy distintos. El mensaje genético propiamente dicho (la información que determina la estructura de las proteínas) está escrito en la conocida clave genética de tripletes. El mensaje es lineal y, su lectura, extrínseca. La sucesión de tripletes de los grupos químicos llamados bases a lo largo de una de las cadenas de la doble hélice determina la sucesión de aminoácidos que se unen para formar una cadena de proteína. No parece haber una relación estructural innata entre cada triplete de bases y el aminoácido por él determinado. Por el contrario, en la transferencia de información interviene una compleja maquinaria extrínseca al ADN: ARN mensajero, ribosomas, ARN de transferencia y una batería de enzimas auxiliares.

El ADN no sólo contiene el mensaje genético, sino también instrucciones para su lectura selectiva. Lo poco que se sabe sobre este tipo de información procede de estudios de la regulación genética en bacterias. Una proteína represora desactiva bloques de genes del cromosoma bacteriano uniéndose a una región del ADN (el operador), cuya peculiar secuencia de bases es reconocida por aquélla; los genes se activan cuando ciertas moléculas pequeñas se unen al represor y le hacen separarse del operador. Aunque queda mucho por averiguar, la mayoría de los biólogos moleculares espera encontrar mecanismos similares en la regulación genética de los organismos superiores a las bacterias. Un represor bacteriano reconoce directamente su operador, probablemente porque se forman puentes de hidrógeno entre los átomos de nitrógeno y oxígeno de sus aminoácidos y los de los bordes libres de los pares de bases de la doble hélice del ADN. Una conformación particular de la proteína se ajusta perfectamente sobre la hélice. La información que determina la regulación genética es, por tanto, tridimensional y su interpretación es intrínseca: depende de la estructura de la proteína y la hélice.

¿Desempeña el ADN sólo un papel pasivo en el reconocimiento, insertándose las cadenas laterales de los aminoácidos en los surcos de una doble hélice estática, o bien la secuencia de bases modifica la estructura de la hélice, contribuyendo a su reconocimiento por las proteínas reguladoras? Hace sólo cuatro años la pregunta hubiera sido irrelevante, porque, dados los conocimientos de que se disponía, era totalmente imposible de contestar.

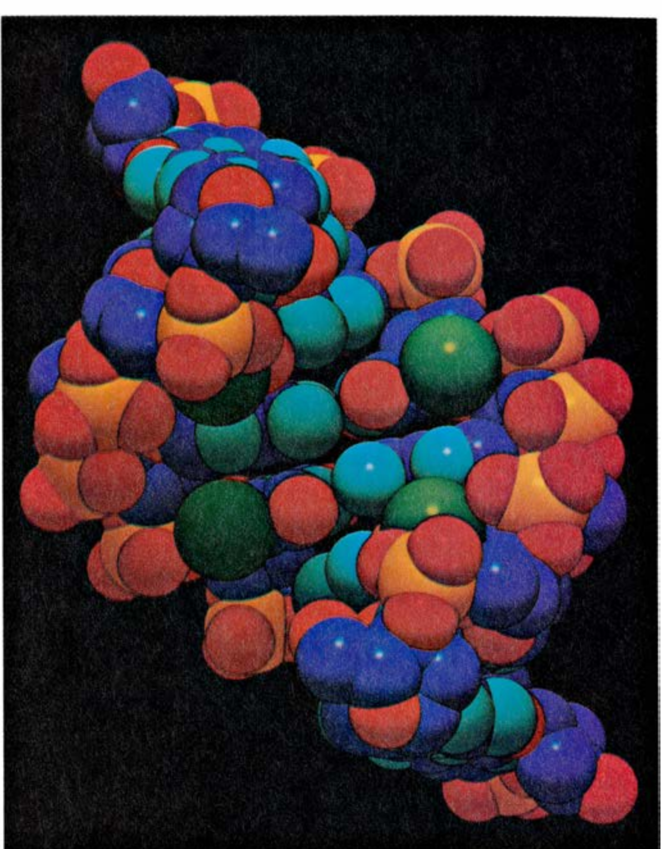
James D. Watson y Francis Crick propusieron su estructura bihelicoidal del ADN en 1953, basándose en fotografías de difracción de rayos X tomadas por Rosalind Franklin y Maurice Wilkins. Se trataba de fotografías de fibras de ADN. Los estudios posteriores de tales difracciones condujeron a la idea de que las fibras del ADN pueden existir en dos formas: ADN *B*, cuando la humedad es alta, y ADN *A*, si la humedad es menor. Se construyeron modelos moleculares para ambas formas. La cantidad de información estructural que se puede extraer de los modelos de difracción de las fibras está limitada inevitablemente por el desorden de las cadenas de ADN respecto del eje de las fibras. A lo sumo podía conseguirse una hélice promedio general. Las variaciones locales que pudieran inducir las distintas secuencias de bases no se detectaban.

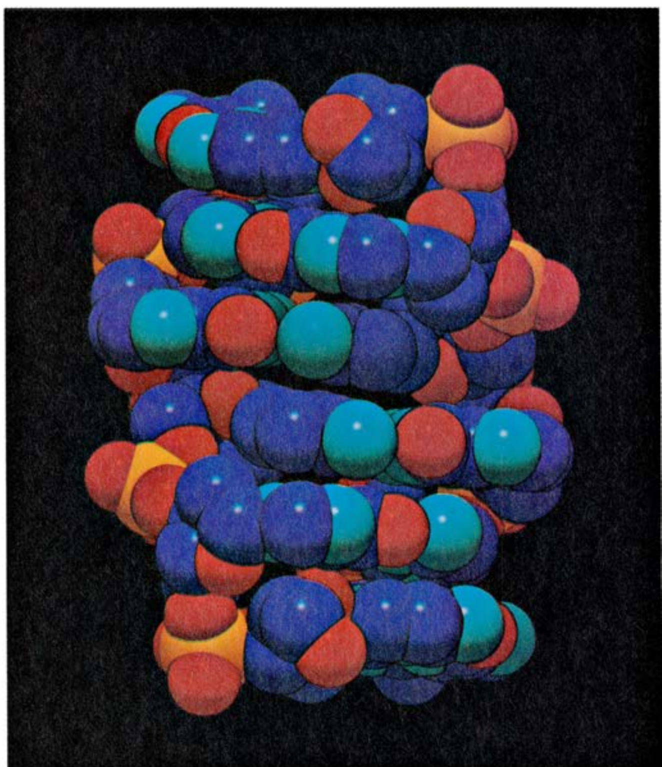
Hoy pueden ya detectarse y medirse con precisión. Los métodos actuales para la síntesis orgánica de pequeñas moléculas de ADN de cualquier secuencia deseada han permitido la síntesis de oligómeros (cadenas cortas) de ADN de cuatro a 24 bases en cantidad y pureza suficientes para cristalizarlas y para estudiar la difracción de rayos X por un solo cristal. Este artículo pretende ser un escaparate de algunos de los nuevos e interesantes resultados que empieza a producir el análisis estructural de cortas moléculas bicatenarias de ADN

Difracción de fibras

La difracción de rayos X de fibras tensas y capas finas de ADN natural reveló la estructura básica de los dos tipos fundamentales de doble hélice: la configuración *B*, que es estable cuando la humedad relativa es de alrededor del 92 por ciento, y la configuración *A*, que adquieren la mayoría de las secuencias cuando la humedad baja al 75 por ciento, aproximadamente. Ambas tienen la forma de una escalera flexible de ADN, enrollada helicoidalmente alrededor de un eje central. Los dos largueros de la cadena son cadenas en las que alternan los anillos del azúcar desoxirribosa y grupos fosfato. Los travesaños son pares de bases, una purina y una pirimidina. Hay dos purinas diferentes, cada una con dos anillos, la adenina (*A*) y la guanina (*G*), y dos pirimidinas, cada una con un solo anillo, la timina (*T*) y la citosina (*C*). Los pares de bases están unidos por puentes de hidrógeno. La adenina se une normalmente con la timina por dos puentes de hidrógeno y la guanina con la citosina por tres.

Los dos anclajes de cada par de bases a los anillos de azúcar no quedan diametralmente opuestos en los extremos de bases [*véase la figura 4*], lo que resulta importante para la geometría de la doble hélice. El borde cóncavo del par de bases, donde están los anclajes, se llama borde del surco menor y, el borde convexo, borde del surco mayor. Cuando se apilan los pares de bases para formar la hélice, las cadenas fosfatadas constituyen los lados de un surco mayor y un surco menor, enrollados al-





1. MODELOS DE TRES TIPOS DE DOBLE HELICE, según imágenes generadas en el ordenador por Nelson Max, del Laboratorio Nacional Lawrence Livermore de los Estados Unidos. Muestran tres de las cortas moléculas bihelicoidales cuyas estructuras han sido resueltas por análisis de difracción de rayos X en cristales aislados. Los átomos de carbono aparecen en azul oscuro, los de nitrógeno en azul claro, los de oxígeno en rojo y los de fósforo en amarillo. Los átomos de bromo añadidos a algunas bases están en verde. El ADN *B* (*izquierda*) está representado por una molécula de doce pares de bases cuya estructura resolvieron el autor y sus colaboradores. En el centro se aprecia el surco mayor de la hélice. El esqueleto de fosfatos resalta el giro dextrorso. La molécula de ocho bases de ADN *A* (*arriba a la derecha*) fue resuelta por Olga Kennard, Zippora Shakked y M. A. Viswamitra. También es dextrorsa. La representación, que apunta directamente al profundo surco mayor, muestra claramente la inclinación de los pares de bases respecto al eje vertical de la hélice. La molécula de ADN *Z* (*abajo a la derecha*) fue establecida por Andrew H.-J. Wang, Alexander Rich y sus colaboradores. Como indican los fosfatos situados arriba a la derecha y abajo a la izquierda, la hélice *Z* es sinistrorsa. Las cadenas zigzaguean y los pares de bases se apilan de dos en dos.

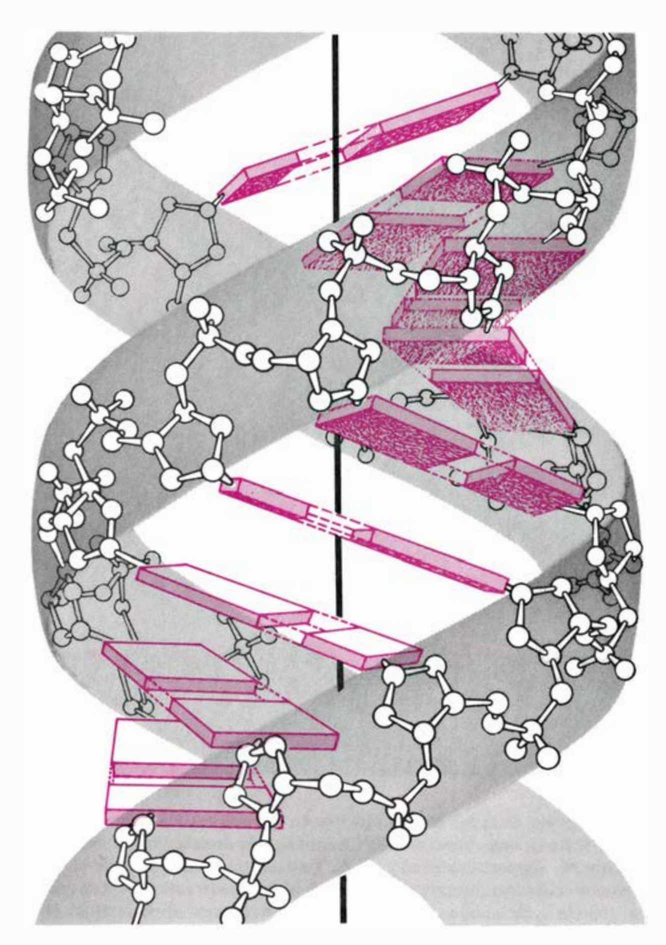
rededor de la hélice, y los bordes de las bases son los fondos de los surcos.

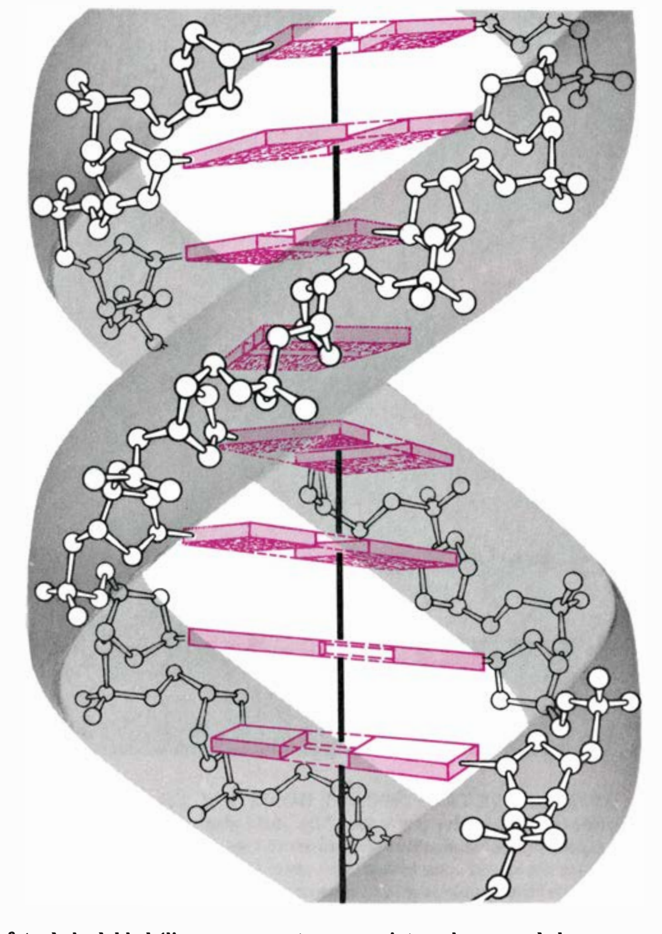
El fondo del surco mayor contiene átomos de nitrógeno y oxígeno que pueden formar puentes de hidrógeno con las cadenas laterales de los aminoácidos de una proteína y, por tanto, pueden contener una información intrínseca valiosa. La distribución de estos átomos depende del par de bases de que se trate. Yendo de la purina a la pirimidina, un par *A-T* ofrece un átomo de nitrógeno (aceptor de hidrógeno), un grupo NH_2 (dador) y un átomo de oxígeno (otro aceptor). Por el contrario, el par *G-C* ofrece los mismos grupos en orden diferente: primero un nitrógeno (un aceptor), luego un oxígeno (un aceptor) y finalmente un NH_2 (un dador). Como cada par de bases (*A-T* y *G-C*) también puede darse la vuelta (para convertirse en *T-A* y *C-G*), un represor u otra proteína reguladora se encuentra con cuatro distribuciones posibles en cada travesaño de la hélice. Así, el fondo del surco mayor contiene un mensaje, la secuencia de bases del ADN, en una forma que pueden leer directamente otras macromoléculas.

El surco menor es menos informativo: los pares *A-T* y *T-A* contienen solamente dos grupos aceptores, y los pares *G-C* y *C-G* difieren de aquéllos sólo por la inclusión de un dador, NH_2, entre los aceptores. Por eso el surco menor es menos apropiado para la lectura de la información. Como veremos luego, el surco menor tiene otra función importante en el ADN *B*.

Los ADN *A* y *B* difieren principalmente en la posición y la inclinación del par de bases respecto del eje de la hélice: en el ADN *B* la inclinación es próxima a cero (los pares de bases están casi perpendiculares al eje de la hélice) y el eje atraviesa aproximadamente los pares de bases por su centro. Aunque el surco menor es más estrecho que el mayor, ambos son casi igual de profundos: los bordes de los pares de bases quedan aproximadamente a la misma distancia de la superficie de un cilindro que contuviera a la doble hélice. En el ADN *A*, por el contrario, los pares de bases están inclinados de 13 a 19 grados respecto de la perpendicular. Además, están desplazados hacia el exterior de la hélice, cuyo eje queda en el surco mayor y no toca las bases. Por consiguiente, el surco menor es poco profundo, algo más que una depresión helicoidal en el exterior del cilindro. El surco mayor es profundo como una caverna; se extiende desde la superficie hasta el eje central de la hélice e incluye parte del espacio hacia el lado opuesto.

El ADN *B* tiene, por término medio, 10 pares de bases por vuelta de la hélice y un espaciamiento entre pares de bases contiguas de 0,34 nanómetros a lo largo del eje de la hélice. La hélice *A* está más próxima a 11 pares de bases por vuelta y el espaciamiento entre pares de bases a lo largo del eje es de sólo 0,29 nanómetros. Estas distancias son promedios obtenidos con fibras; los estudios con cristales revelaron después variaciones locales importantes respecto de los valores medios. (El modelo original de Watson y Crick tenía 10 pares de bases por vuelta y el apilamiento perpendicular de bases característico del ADN *B*, pero las bases estaban desplazadas del centro para dejar un hueco central, más típico de la forma *A*. La posición ocupada por los anillos de azúcar en el modelo original





2. ESQUEMAS de ADN *A* (*izquierda*) y ADN *B* (*derecha*), por Irving Geiss, basados en los estudios de fibras de ADN de Struther Arnott y otros, anteriores a los estudios de cristales tratados en este trabajo. Las cadenas de azúcar y fosfato de la doble hélice se representan como cintas y los pares de bases que las conectan, como travesaños. Los pares de bases del ADN *A* están inclinados y evitan al eje de la doble hélice. En el ADN *B*, el eje de la hélice los corta.

era también más parecida a la de la forma *A*.)

Tal era el estado aproximado del conocimiento sobre la estructura del ADN bicatenario durante los 25 años que siguieron al descubrimiento de Watson y Crick. Muchas preguntas quedaban por contestar, y otras tantas ni siquiera tenían respuesta. ¿Por qué prevalece la hélice *B* con mucha humedad? ¿Qué forma domina en el ADN cromosómico? ¿Puede convertirse una forma en otra en un organismo vivo? ¿Leen realmente los represores y otras proteínas la información del fondo de los surcos cuando se unen a secuencias específicas del ADN?

A finales de la década de 1970, cuando los progresos de la síntesis de ADN permitieron por primera vez estudiar cristales aislados de moléculas cortas de cualquier secuencia deseada, el grupo de Alexander Rich, en el Instituto de Tecnología de Massachusetts, y el nuestro, que estaba entonces en el Instituto de Tecnología de California, emprendimos independientemente la síntesis de copolímeros *C-G*: breves moléculas compuestas exclusivamente de citosinas y guaninas alternadas. Había una razón para elegir esta secuencia: las investigaciones con soluciones de ADN sugerían que el copolímero *C-G* sufría una transición estructural en presencia de mucha sal o mucho alcohol y era, por tanto, el material más apropiado para estudiar transiciones entre tipos de hélice.

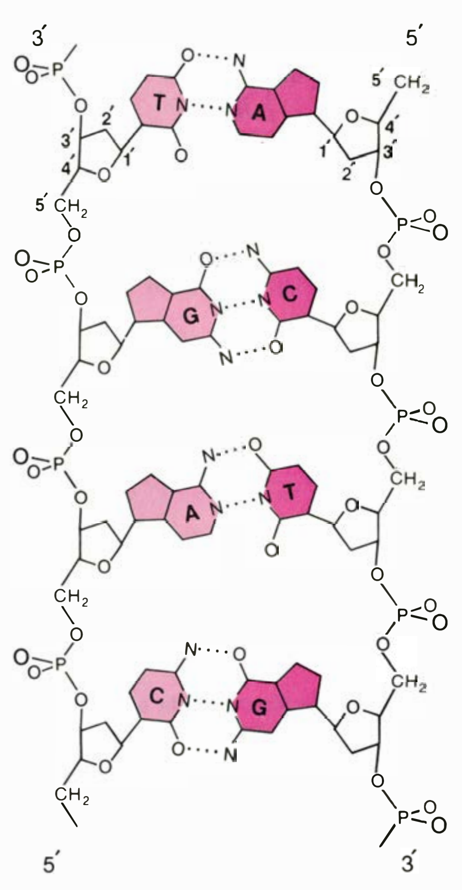
Estudios con cristales aislados

Andrew H.-J. Wang, del Instituto de Tecnología de Massachusetts (MIT), resolvió la estructura del hexámero *CGCGCG* a mediados de 1979 y Horace R. Drew la del tetrámero *CGCG* unos pocos meses después en nuestro laboratorio. Para sorpresa de todos, las dos moléculas resultaron no ser ADN *A* ni *B*. Ni siquiera eran hélices dextrorsas. Eran hélices sinistrorsas de un tipo completamente nuevo, con unas cadenas zigzagueantes peculiares que dieron lugar al nombre de hélice *Z*. Los estudios con fibras nos habían instado a preguntar: ¿tienen estas moléculas la forma *A* o la *B*? La naturaleza nos castigó con una respuesta rotunda: no.

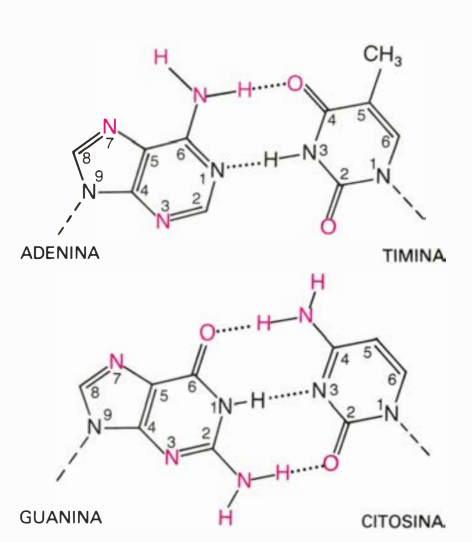
Aunque las dos primeras moléculas de ADN bicatenario cuya estructura se resolvió por difracción de rayos X en cristales aislados rompieron todos los pronósticos, las dos siguientes nos dieron un ejemplo de ADN *B* y otro de ADN *A*. Drew sintetizó el dodecámero *CGCGAATTCGCG*, polímero que elegimos por dos razones. Contiene extremos *CGCG*, compatibles con una estructura *Z*, y un centro *AATT*, incompatible con ella, lo que permite averiguar si *CGCG* es capaz de adoptar la forma *Z* en un entorno extraño. Incluye también la secuencia *GAATTC*, reconocida y rota por *Eco* *R*I, una de las endonucleasas de restricción, enzimas que cortan las moléculas de ADN por sitios específicos. (El lector habrá advertido que el oligómero artificial *CGCGAATTCGCG*, como otros sintetizados para estos estudios, es complementario de sí mismo: la secuencia leída de izquierda a derecha es complementaria de la leída de derecha a izquierda, de modo que dos cadenas cualesquiera pueden combinarse para formar una doble hélice. Así se ahorra mucho tiempo y esfuerzo, porque sólo hay que sintetizar una de las cadenas de la doble hélice.) La estructura de esa molécula resultó no tener nada que ver con el ADN *Z*; es un ejemplo clásico de hélice *B*. Benjamin N. Conner sintetizó también, en nuestro laboratorio, el tetrámero *CCGG* y demostró que era un fragmento de ADN *A*.

Características estructurales

Varios laboratorios publicaron poco después la estructura de otras moléculas bihelicoidales de ADN. La forma *A* está representada por *CCGG* y *GGCCGGCC*, analizadas por el grupo del MIT, y por *GGTATACC*, de un grupo que incluye a Olga Kennard, de la Universidad de Cambridge, y a Zippora Shakked, del Instituto Weizmann de Ciencias de Israel. Wang ha resuelto también la estructura de una forma *A* que mezcla ARN y ADN: (*GCG*)*TATACGC*. (Las letras entre paréntesis representan bases de ARN, de modo que los tres pares de bases de los extremos son híbridos de ADN y ARN y los cuatro centrales son ADN puro.) El ADN *B* está representado por ahora sólo por *CGCGAATTCGCG* y sus derivados. El ADN *Z* se encuentra en los ya mencionados *CGCG* y *CGCGCG* y una variante de este último, en el que se añaden grupos metilos (CH_3) a las citosinas, una modificación que favorece la estructura *Z*. Podemos analizar estas moléculas para ver si cumplen las predicciones sobre ADN *A* y *B* de los estudios de difracción de fibras, cuánto se diferencian de las propiedades medias de las hélices y hasta qué punto esas diferencias pueden atribuirse a la secuencia de bases del ADN.



3. HELICE DESENROLLADA para indicar las cadenas de azúcar y fosfato y los travesaños de pares de bases. Las cadenas tienen direcciones opuestas; sus extremos 5′ y 3′ se llaman así por la orientación de los átomos de carbono 5′ y 3′ de los azúcares. Cada par tiene una purina, adenina (*A*) o guanina (*G*), y una pirimidina, timina (*T*) o citosina (*C*), conectadas por puentes de hidrógeno.



4. PARES DE BASES en detalle. Los trazos laterales indican los enlaces a los azúcares de las cadenas. Cada par de bases tiene un borde al surco mayor (arriba) y otro al surco menor (abajo). Los grupos amino (NH_2) son dadores potenciales de hidrógeno, capaces de formar puentes de hidrógeno con represores y otros reguladores; los átomos de nitrógeno y oxígeno indicados en color son aceptores potenciales de hidrógeno. Las proteínas reguladoras probablemente "lean" la distribución característica de dadores y aceptores en cada par.

	ADN *A*	ADN *B*	ADN *Z*
GIRO	DEXTRORSO	DEXTRORSO	SINISTRORSO
ENROLLAMIENTO (GRADOS) MEDIA Y DESVIACION TIPICA INTERVALO OBSERVADO	33,1 ± 5,9 16,1 a 44,1	35,9 ± 4,3 27,7 a 42,0	*G-C*: −51,3 ± 1,6 *C-G*: −8,5 ± 1,1
PARES DE BASES POR VUELTA	10,9	10,0	12,0
ELEVACION DE LA HELICE POR CADA PAR DE BASES (NANOMETROS)	0,292 ± 0,039	0,336 ± 0,042	*G-C*: 0,352 ± 0,022 *C-G*: 0,413 ± 0,018
INCLINACION DE LAS BASES (GRADOS)	13,0 ± 1,9	−2,0 ± 4,6	8,8 ± 0,7
ALABEO (GRADOS)	15,4 ± 6,2	11,7 ± 4,8	4,4 ± 2,8
BALANCEO (GRADOS)	5,9 ± 4,7	−1,0 ± 5,5	3,4 ± 2,1

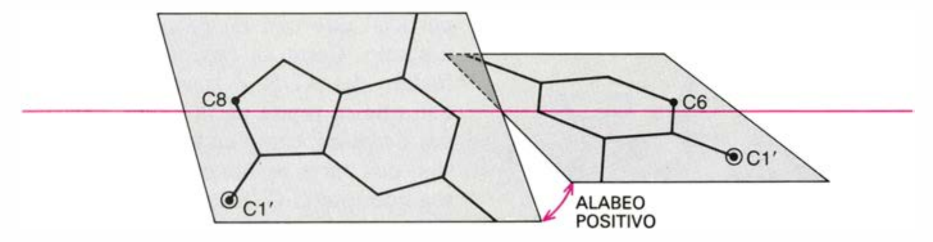
5. PARAMETROS HELICOIDALES MEDIOS y sus desviaciones típicas en tres tipos de ADN. Los datos se basan en estudios con rayos X de cristales aislados de moléculas mencionadas en este artículo.

A partir de los estudios con cristales aislados se han determinado valores medios de las características estructurales de cada clase de hélice [*véase la figura 5*]. Los ADN *A* y *B* son dextrorsos, con un enrollamiento medio de una base a la siguiente de 33,1 y 35,9 grados, respectivamente, que corresponden a 10,9 y 10,0 pares de bases por vuelta completa, en estrecho acuerdo con las predicciones de los estudios con fibras. El espaciamiento de los pares de bases a lo largo de la hélice tampoco proporciona sorpresas. Lo llamativo es el grado de variación del enrollamiento respecto de los valores medios: la variación típica es de ±6 grados para el ADN *A* y de ±4 grados para la forma *B*. El enrollamiento entre una base y la siguiente va de 16 a 44 grados en el ADN *A* y de 28 a 42 grados en el ADN *B*. Si se conoce la secuencia de bases puede predecirse esta variación en el enrollamiento de la hélice por el método que describiré después. (¿Constituye esta variación una señal por la que el represor y otras moléculas reguladoras reconocen una secuencia específica de bases? La idea es atractiva.) Las bases se desvían de la perpendicularidad al eje de la hélice aproximadamente lo que ya se esperaba: 13 grados en el ADN *A* y unos 2 en el ADN *B*.

Los pronósticos basados en los estudios con fibras son menos acertados respecto al alabeo de pares de bases. Este alabeo es una rotación alrededor del eje mayor del par de cada base respecto de su compañera de par: se dice que es positivo cuando, mirando a lo largo del eje, la base más próxima ha girado en el sentido de las agujas del reloj [*véase la figura 6*]. Los estudios con fibras se han ocupado poco del alabeo. Una tabla de coordenadas del ADN *A* publicada en 1972 implicaba un alabeo negativo y una revisión de 1981 daba el signo apropiado, pero con una magnitud de sólo 8 grados.

Los estudios de cristales muestran que el alabeo es siempre positivo, con valor medio de 15 grados en el ADN *A* y de 12 en el *B*. Los valores concretos van de los 3 a los 25 grados. En hélices dextrorsas el alabeo positivo mejora el apilamiento de las bases de cada cadena, aumentando el solapamiento entre una base y sus contiguas. Pero el alabeo positivo acerca demasiado las purinas (*G* y *A*) situadas en pares de bases contiguos y cadenas opuestas. Esto altera, como ha sugerido C. R. Calladine, de Cambridge, varios parámetros importantes de la estructura helicoidal, dependientes de la secuencia de bases, entre ellos el propio alabeo, el enrollamiento helicoidal y el llamado balanceo de los pares de bases.

El balanceo del par de bases indica la rotación del par de bases (es decir, del mejor plano medio de la purina y la pirimidina) alrededor del eje mayor del par. Si dos pares de bases sucesivos de la hélice rotan en sentido opuesto, abren un ángulo entre ellas, con entrada por el surco mayor o por el surco menor. El balanceo entre dos pares de bases se define como positivo si el ángulo se abre hacia el surco menor y como negativo si se abre hacia el surco mayor. Los estudios con fibras nos permiten medir estos ángulos, pero los estudios de cristales demuestran que el balanceo es uno de los parámetros importantes de la hélice. Como en el ADN *B*, los pares de bases son casi perpendiculares al eje de la hélice, no hay un balanceo sistemático de cada par de bases respecto al siguiente. En la forma *A*, por el contrario, la escalera del ADN se enrolla alrededor del eje de la hélice de modo que los planos de las



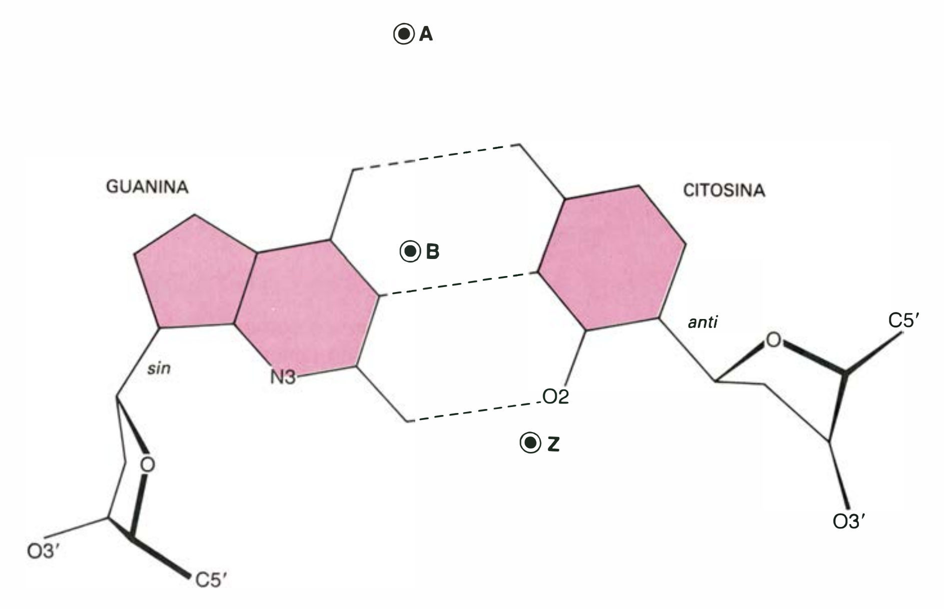
6. DEFINICION DE ALABEO de un par de bases en un dibujo esquemático de un par purina-pirimidina. El alabeo se considera positivo si la base más próxima, vista a lo largo del eje mayor (*color*), ha girado en el sentido de las agujas del reloj. Nótese que el alabeo desplaza los átomos C1′, por los que las bases se unen a la cadena de azúcar y fosfato. En los tres tipos de hélice el alabeo es siempre positivo.

bases se abren, como los segmentos de un acordeón, hacia el surco menor, con un balanceo medio de 6 grados. Aún más significativo es que tanto la hélice *A* como la *B* presentan una variación de ±5 grados respecto de los valores medios.

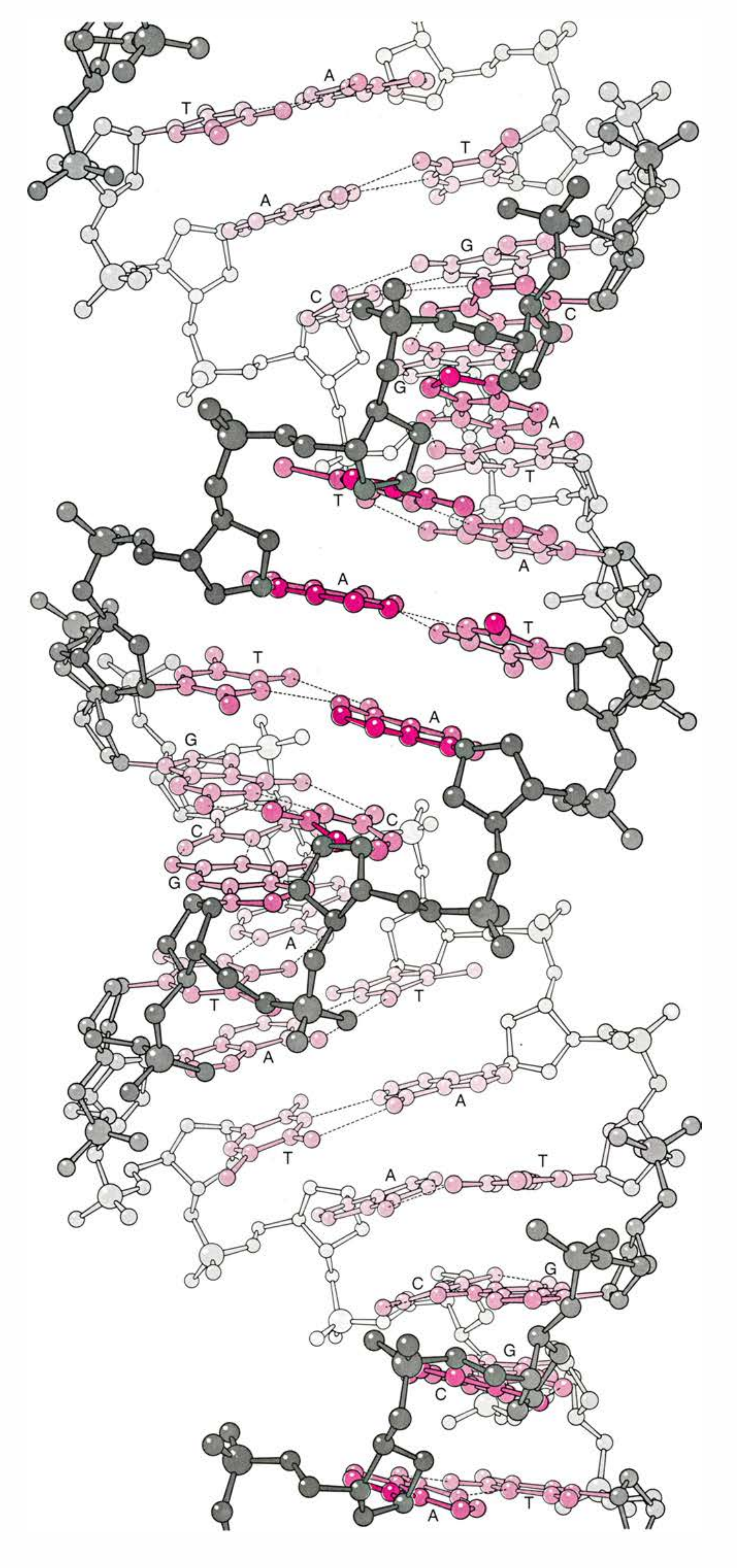
Extrapolación a hélices infinitas

La molécula más larga de ADN estudiada por difracción de rayos X en cristales aislados es el dodecámero *CGCGAATTCGCG*. Se puede tener una idea del aspecto de hélices más largas prolongando tres moléculas típicas [*véase la figura 1*] para producir hélices largas [*véanse las figura 8 a 13*]. Ello se consigue programando un ordenador para que gire la imagen de una molécula helicoidal dada alrededor de su eje hasta que los átomos del comienzo de la hélice girada coincidan con los átomos homólogos del final de la imagen original, y repitiendo el proceso cuantas veces se desee. De esta manera, por ejemplo, se han repetido cuatro veces los seis pares de bases centrales de la hélice *A* de secuencia *GGTATACC* para producir la hélice de 24 pares de bases *GTATACGTATACGTATACG-TATAC*. Así se ha podido generar por primera vez una larga doble hélice a partir de datos obtenidos con rayos X en cristales aislados, sin promediar ni redondear las fluctuaciones locales inducidas por la secuencia de bases en la estructura de la hélice.

Como dejaban esperar los estudios con fibras, la hélice *A* es corta y gorda, con un profundo surco mayor y un surco menor ancho y superficial. La hélice *B* es más delgada (y más larga para el mismo número de pares de bases), con un surco mayor ancho y un surco menor estrecho de profundidades parecidas. Por el contrario, la nueva hélice sinistrorsa *Z* es muy delgada y alargada, con un surco menor estrecho y profundo y un "surco" mayor tan su-



7. DOS CONFORMACIONES del enlace C-N que conecta cada base a su azúcar. La conformación *anti* (*derecha*) se da en todo el ADN *A* y *B* y en las citosinas del ADN *Z*; la conformación *sin* (*izquierda*) aparece en las guaninas del ADN *Z*. La alternancia de una purina y de una pirimidina (*G, C*) en cada cadena del ADN *Z* permite la alternancia *sin-anti*, que provoca el zigzagueo de las cadenas. Los tres puntos marcados *A*, *B* y *Z* indican la situación del eje de la hélice respecto de un par de bases en cada tipo de hélice. El eje cae en el surco mayor en el ADN *A*, atraviesa el par de bases en el *B* y cae en el surco menor en el *Z*.



perficial que casi no parece que haya ninguno. Como la cadena de azúcar y fosfato del ADN *Z* zigzaguea, la unidad que se repite en la hélice no es el par de bases, como en los ADN *A* y *B*, sino dos pares de bases sucesivos: *G-C* seguidos por *C-G*. Ello se debe al modo diferente en que las citosinas y las guaninas se unen a los anillos de azúcar del ADN *Z*.

En las citosinas el azúcar aparece girado alrededor de su enlace con la base, de manera que el anillo fruncido del azúcar parece huir del surco menor [*véase la figura 7*]. Esta conformación, llamada *anti*, se encuentra también en todas las bases de los ADN *A* y *B*. En el ADN *Z*, por el contrario, el anillo del azúcar de cada guanina está girado 180 grados, es decir, doblado hacia el surco menor. Esta configuración *sin* es factible estereoquímicamente cuando el azúcar se une al anillo pequeño, pentagonal, de una purina, en vez de al anillo hexagonal de una pirimidina: el azúcar en configuración *sin* quedaría demasiado cerca del átomo de oxígeno de la posición O2 de la citosina o la timina. La alternancia de una purina (*G*) y una pirimidina (*C*) en cada cadena del ADN *Z* permite una alternancia *sin-anti* en los enlaces entre los azúcares y las bases y produce la cadena zigzagueante. El curso de la cadena entre los dos átomos de fósforo que flanquean una guanina queda casi paralelo al eje de la hélice; el curso entre los dos átomos de fósforo que flanquean una citosina, por el contrario, es tangencial a un cilindro que envolviera a la hélice y perpendicular a su eje. Por eso la unidad que se repite a lo largo de la hélice es de dos pares de bases; aunque la hélice *Z* tiene doce pares de bases por vuelta, se puede considerar hélice sinistrorsa con seis elementos de dos pares de bases cada uno por cada vuelta de hélice.

ARN

Sólo hemos comentado hasta aquí el ADN (ácido desoxirribonucleico), portador de la información genética en el núcleo de la célula. La expresión de esa información depende en gran parte de

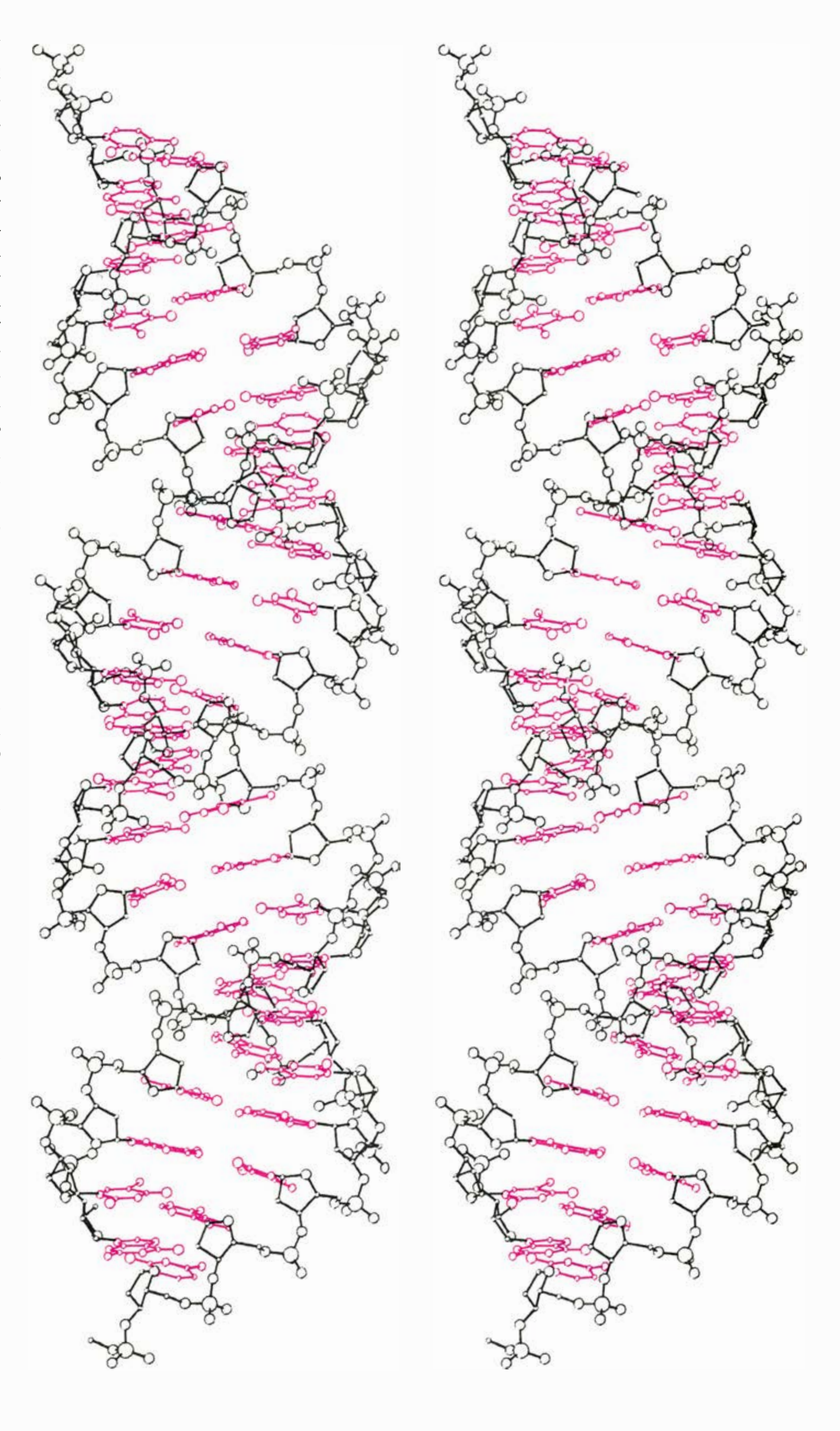
8. HELICE DE ADN *A*, según un esquema basado en el par estereoscópico obtenido en un ordenador y representado en la figura 9. Para generar los dibujos se extrapoló la estructura de las seis bases centrales del octámero *GGTATACC*, determinada por Kennard, Shakked y Viswamitra. Los pares de bases no son figurados, sino que se dedujeron de la estructura del cristal. Los fosfatos de las cadenas opuestas se encaran en el surco mayor.

un ácido nucleico muy similar, el ARN (ácido ribonucleico). La secuencia de bases de una cadena de ADN se transcribe a una sola cadena de ARN mensajero; las moléculas de ARN de transferencia leen la secuencia y traen los aminoácidos correspondientes a los ribosomas para su engarce en una cadena de proteína; los ribosomas contienen también ARN. Hay sólo dos diferencias entre el ADN y el ARN. La timina del ADN se reemplaza en el ARN por uracilo, que carece del grupo metilo lateral de la timina, y, en vez de los anillos de desoxirribosa de la cadena del ADN, la cadena del ARN tiene anillos de ribosa. En la desoxirribosa un átomo de hidrógeno reemplaza al grupo hidroxilo (OH) unido a uno de los átomos de carbono (C2′) del anillo de la ribosa.

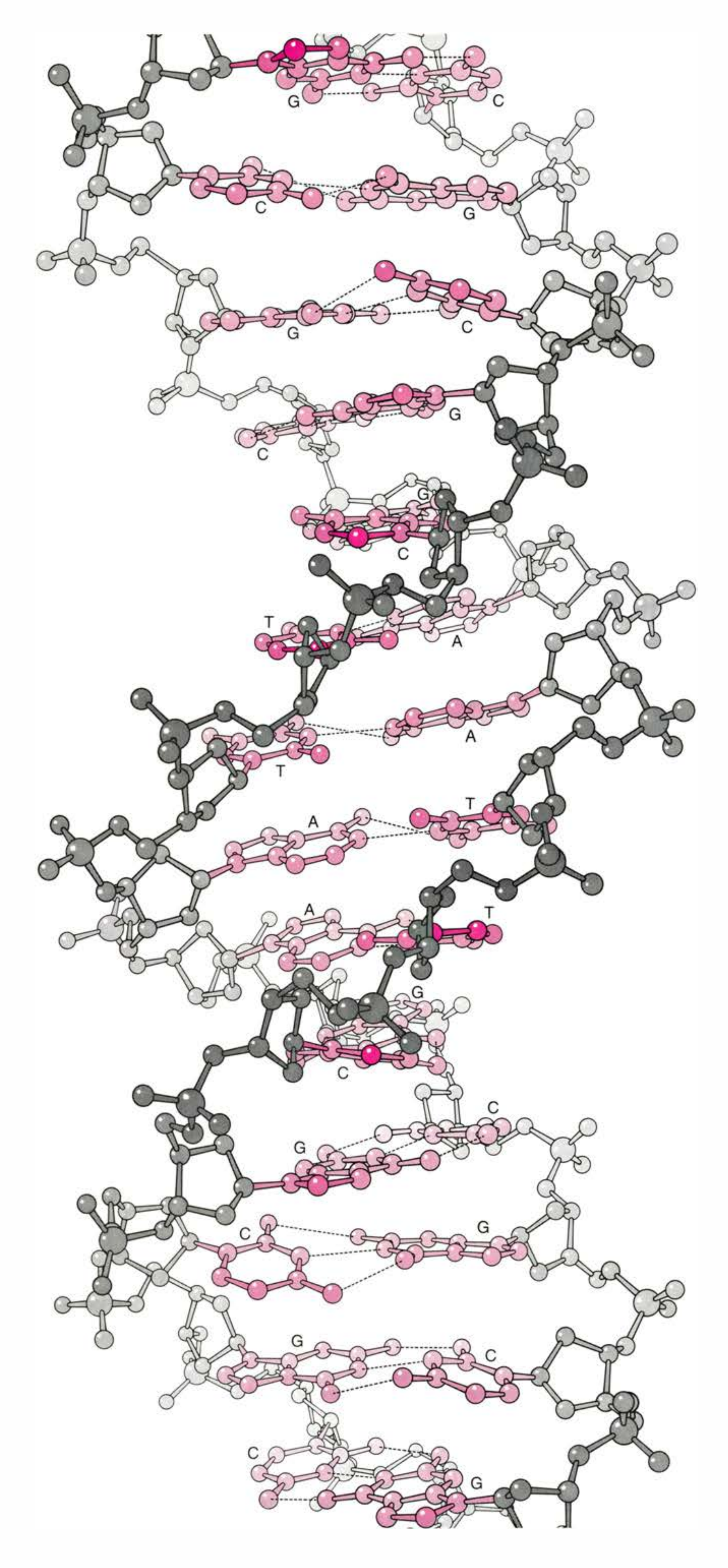
Una consecuencia importante del hidroxilo adicional del ARN es que el ARN no parece poder enrollarse en una doble hélice *B*. Si se supone una estructura *B* para el ARN, el grupo OH quedaría rodeado por los átomos del siguiente grupo fosfato, el anillo del azúcar y la base. El átomo de oxígeno adicional quedaría inaceptablemente próximo a otros átomos, haciendo la estructura estereoquímicamente desfavorable. En el ADN *A*, por el contrario, el grupo 2′OH sobresale de la superficie de la hélice y se aleja de los átomos vecinos.

Los fragmentos bicatenarios formados al enrollarse la molécula de ARN transferente sobre sí misma tienen que ser, por consiguiente, variaciones de la hélice *A*. Si, como se ha sugerido, las exiguas regiones autocomplementarias del ARN mensajero también pueden formar estructuras bicatenarias en horquilla, mantenidas por puentes de hidrógeno, éstas deben ser también ARN *A* y no *B*. (En el único ejemplo conocido de participación de ARN en una hélice *B*, un híbrido sintético con adeninas en la cadena de ARN y timinas en la cadena de ADN, el grupo OH intruso se acomoda formando un puente de hidrógeno con el átomo O4′ de la ribosa siguiente.) Nadie ha sugerido seriamente la posibilidad de un ARN *Z*, pero un examen detallado de su estructura indica que el OH extra del azúcar quedaría probablemente demasiado cerca del átomo O2 que sobresale del anillo de la citosina.

El análisis estructural por rayos X de cristales aislados ha definido tres tipos fundamentales de doble hélice del ADN, en vez de dos, y un tipo de doble hélice de ARN. La forma *B* del ADN es la estable si hay mucha humedad, lo



9. PAR ESTEREOSCOPICO obtenido a partir de las coordenadas tridimensionales del hexámero *GTATAC*, un ADN *A*. Se eliminaron los pares de bases de los extremos del octámero (carecían de grupos fosfato), se giró y trasladó la imagen respecto del eje principal hasta que sus dos fosfatos inferiores coincidieron con los fosfatos superiores de la imagen no desplazada y se repitió el proceso varias veces para conseguir una especie de hélice *A* infinita. Se puede generar así una larga molécula de un ADN específico sin sacrificar las variaciones estructurales locales impuestas por su secuencia. La mejor forma de ver este dibujo en relieve es usar un estereoscopio. Con un poco de práctica se puede ver una imagen tridimensional sin estereoscopio si uno consigue desacoplar dos reflejos que normalmente van asociados: separar la mirada de los ojos como si se mirara un objeto lejano mientras se enfoca un objeto próximo. Nótese la profundidad del surco mayor. La hélice parece dos cintas enrolladas al exterior de un cilindro.



que sugiere la presencia de la hélice *B* en el núcleo de la célula. Si durante la transcripción del ADN a ARN mensajero se forma transitoriamente un híbrido ADN-ARN, es probable que adopte la estructura *A*. En ciertas condiciones muy duras, el ADN dextrorso con purinas y pirimidinas alternantes puede saltar a la forma sinistrorsa *Z*. Para la mayoría de las secuencias y en presencia de agua, sin embargo, lo normal parece ser el ADN *B*. ¿Cómo estabiliza el agua la hélice *B*?

Estabilización por hidratación

Una vez establecida la estructura de una molécula de ADN con rayos X se pueden buscar cuidadosamente las moléculas de solvente (agua) que estén perfectamente ordenadas a su alrededor, es decir, tan fuertemente unidas al ADN que ocupen los mismos lugares en todas las moléculas del cristal. (Las moléculas de agua unidas menos fuertemente, que ocupan sitios diferentes en distintas moléculas del cristal, dan difracciones borrosas escasamente distinguibles del ruido de fondo.) Se ha investigado la presencia de agua alrededor de moléculas de los tres tipos de ADN bicatenario, pero sólo se han publicado resultados detallados para el ADN *B*. En éste se encuentran moléculas de agua en la vecindad de casi todos los átomos que podrían formar puentes de hidrógeno con ellas: los oxígenos libres de los fosfatos, los átomos de nitrógeno y oxígeno de los bordes de los pares de bases y, menos frecuentemente, los oxígenos de la cadena principal. La hélice queda cubierta con una capa de agua de una molécula de profundidad.

Se observa una hidratación más compleja en el surco menor del ADN *B* [*véase la figura 15*]. La combinación del alabeo de cada par de bases y del enrollamiento helicoidal entre un par de bases y el siguiente acerca un oxígeno o un nitrógeno de una cadena a un oxígeno o un nitrógeno situado en el par de bases contiguo pero en la cadena opuesta. Una molécula de agua establece un puente entre los dos aceptores de hidrógeno. Una segunda tanda de moléculas de agua forma puentes entre las moléculas de agua de la primera. En conjunto aparece un cordón zigza-

10. HELICE DE ADN *B* que se obtuvo repitiendo los diez pares de bases centrales del dodecámero formado por *CGCGAATTCGCG*. El dibujo muestra el fuerte alabeo de los pares de bases, que mejora el apilamiento de las bases de las dos cadenas.

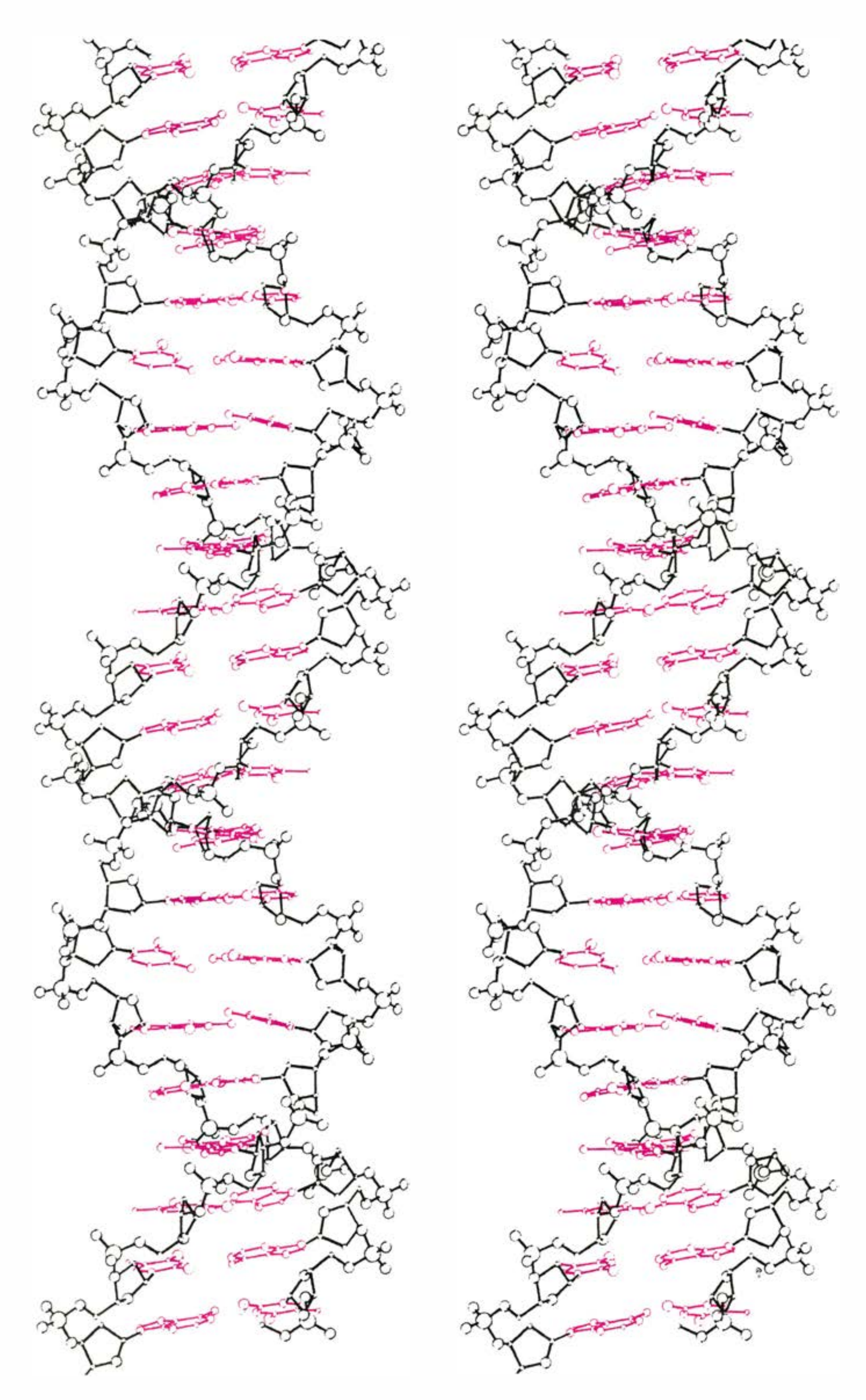
gueante de moléculas de agua en el fondo del surco menor de la hélice *B*. El cordón de moléculas de agua está bien desarrollado en las regiones ricas en pares *A-T*, pero se interrumpe en las regiones ricas en *G-C*, probablemente porque los grupos NH_2 de las guaninas sobresalen en el surco menor. Los pares de bases de ADN sintético en que la guanina haya sido reemplazada por la base modificada inosina, que carece del grupo NH_2 de la guanina, se comportan como pares *T-A* y no como pares *C-G*, por cuanto estabilizan la estructura *B*.

Creemos que este cordón de agua del surco menor estabiliza mucho el ADN *B* y tiene que eliminarse para abrir el surco menor y pueda saltar la hélice a la forma *A*. Un modo de destruir el cordón estabilizante en el laboratorio es el de eliminar las moléculas de agua deshidratando las fibras. Otro, introducir más grupos NH_2 en el surco menor; los estudios de difracción de fibras de Struther Arnott, de la Universidad de Purdue, han demostrado que, al aumentar el contenido *G-C* de la fibra, se torna más fácil la transición de la forma *B* a la forma *A* por secado.

No se ha encontrado una distribución sistemática comparable de moléculas de agua alrededor de las hélices *A* o *Z*. Aunque hay muchas moléculas de agua alrededor de los átomos del ADN capaces de formar puentes de hidrógeno, no existe nada semejante al cordón de hidratación del surco menor del ADN *B* ni nada capaz de mantener su integridad estructural; el amplio surco menor del ADN *A* parece incluso menos hidratado que otras partes de la hélice *A*. Pero no estamos seguros de la hidratación del ADN *A*, porque en los tres cristales estudiados, *CCGG/CCGG* (dos tetrámeros apilados para componer un octámero), *GGCCGGCC* y *GGTATACC*, los pares de bases terminales de cada octámero tocan los surcos menores de los octámeros vecinos, impidiendo parte de la hidratación que cabría esperar si las moléculas estuvieran separadas en solución. A pesar de eso, se comprende el predominio de la forma *B* en condiciones acuosas y su transición a la forma *A* al deshidratarse.

Regulación genética

Consideremos ahora las implicaciones biológicas posiblemente más significativas de los estudios de cristales de ADN con rayos X: la influencia de la secuencia de bases en la estructura local de la hélice y su posible papel en la expresión de la información genética. Si la hélice *Z* sólo se da realmente en unas pocas secuencias con purinas y pirimidinas alternantes, es improbable que tenga mucha importancia en las regiones de los genes que determinan las proteínas. Pero, como han sugerido Rich y otros, el cambio de sentido del enrollamiento helicoidal de ciertas regiones reguladoras del ADN podría ser crucial para la accesibilidad de la información genética y su lectura; podría producirse un cambio en el grado de enrollamiento superhelicoidal del ADN circular.



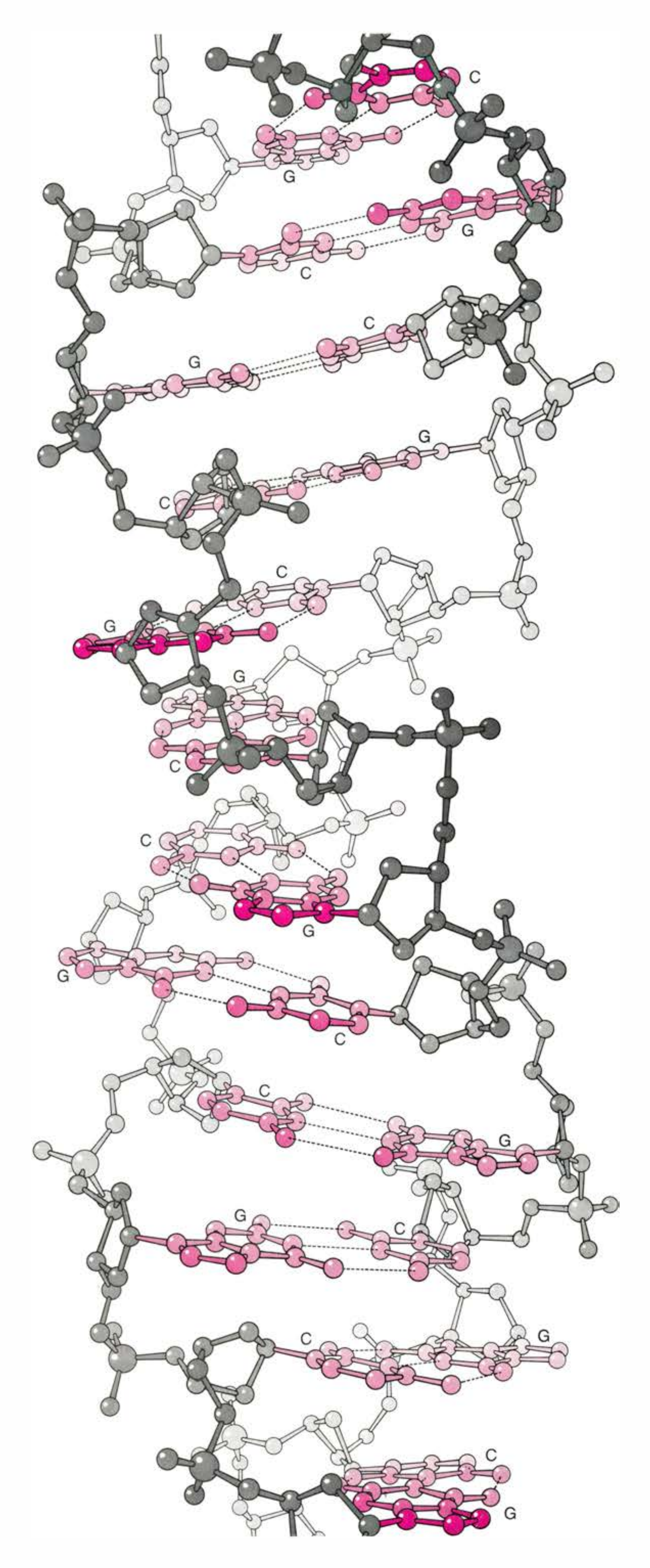
11. PAR ESTEREOSCOPICO que se obtuvo eliminando las bases extremas del dodecámero y repitiendo la secuencia *GCGAATTCGC*. Se observa que en la hélice *B* los surcos mayor y menor tienen casi la misma profundidad, aunque el surco menor es más estrecho. Resulta particularmente estrecho en la región *AATT* (*delante*) y más ancho en las regiones ricas en pares *C-G* (*atrás*). Ilustración del propio autor.

Tal mecanismo constituiría un ejemplo de la influencia de la secuencia de bases sobre la expresión, y no sobre el contenido del mensaje genético. La expresión y la represión selectiva de la información en determinadas células en ciertos momentos, como quiera que se realice, encierra la clave del desarrollo de los organismos pluricelulares. Más aún, la desregulación de la información genética podría ser una de las causas del cáncer y tal vez tenga que ver también con el envejecimiento. Nuestra comprensión de la regulación genética en los organismos superiores está aproximadamente al mismo nivel que la compresión de la clave genética en 1953, cuando Watson y Crick propusieron su doble hélice.

El cambio de sentido del enrollamiento entre las formas *B* y *Z* es un gran cambio conformacional dependiente de la secuencia. El estudio detallado de la estructura local de la hélice en las formas *A* y *B* revela otro tipo de influencia de la estructura sobre la lectura de la información genética. Como indiqué, los valores medios de los parámetros de las hélices suelen ser los que se esperaban a partir de los estudios con fibras; sorprendieron las grandes desviaciones locales.

De la secuencia a la estructura

En el ADN *A* el ángulo medio entre dos pares de bases sucesivos es de 33,1 grados, lo que corresponde casi exactamente a 11 pares de bases por vuelta; pero, ¿quién hubiera predicho que los valores reales variarían entre 16 y 44 grados? El balanceo medio de los pares de bases, 6 grados para el ADN *A* y casi cero para el ADN *B*, podría haberse deducido de la situación de los pares de bases con respecto al eje de la hélice, pero nadie esperaba desviaciones típicas respecto de esas medias de unos ±5 grados en cada caso. Tales fluctuaciones locales indujeron inicialmente a los cristalógrafos a comprobar cuidadosamente sus análisis para asegurarse de que las variaciones observadas no eran meros artefactos de los métodos. Eliminada esta objeción y encontradas variaciones similares en estructuras calculadas de tres maneras distintas, la cuestión obvia fue la de si las fluctuaciones eran predecibles.



12. HELICE DE ADN *Z*, sinistrorsa, de guaninas y citosinas alternantes, generada a partir de los cuatro pares de bases centrales del *CGCGCG* estudiado por Wang, Rich y sus colegas. Los pares de bases aparecen menos alabeados que en las dos hélices dextrorsas. Los fosfatos de las cadenas opuestas ocupan los lados del profundo surco menor.

Se ha intentado repetidas veces explicar las variaciones estructurales por la secuencia de bases del ADN. El intento más productivo, con mucho, fue el de Calladine, que aplicó los principios de la mecánica de fibras elásticas a la doble hélice. Según Calladine, las variaciones locales del enrollamiento y del balanceo de los pares de bases se deben a la incompatibilidad estérica entre las grandes bases púricas, consecuencia del alabeo de los pares de bases (recuérdese que el alabeo apila mejor las bases de cada cadena de la doble hélice).

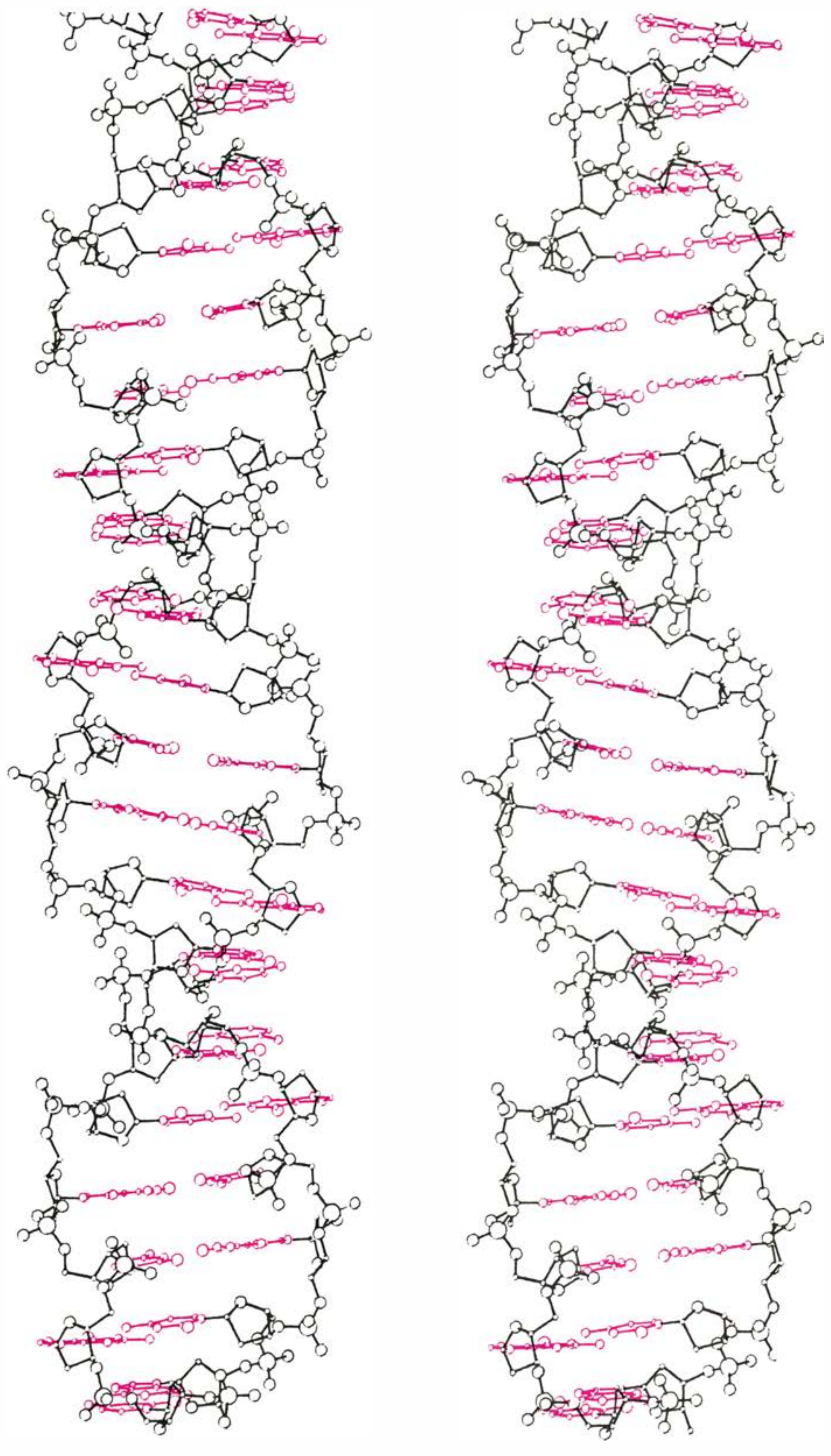
Puesto que las purinas (*G* o *A*) son mayores que las pirimidinas (*C* o *T*), ocupan más de la mitad de cada par de bases. Si dos pares de bases contiguos tienen purinas en cadenas opuestas, estas purinas, vistas a lo largo del eje de la hélice, se solapan ligeramente. Si ambos pares de bases tienen un alabeo positivo, los bordes de las purinas quedarán inaceptablemente próximos y se originará un choque estereoquímico. Si la secuencia de una cadena, en la dirección convencional 5′-3′, es purina-pirimidina, el choque ocurrirá en el surco principal; si es pirimidina-purina, en el surco menor [*véase la figura 16*]. ¿Cómo se pueden evitar estos choques entre los bordes de las purinas?

Una posibilidad es, simplemente, balancear los dos pares de bases, abriendo el ángulo que forman en el lado en que chocan. La respuesta apropiada al choque en el surco menor en la secuencia pirimidina-purina sería separar los bordes que dan al surco menor, haciendo más positivo el ángulo de balanceo. En las secuencias purina-pirimidina, los bordes a separar quedan en el surco mayor y para ello el ángulo de balanceo debe ser más negativo. Drew y yo observamos esta variación del balanceo en función de la secuencia en la estructura *CGCGAATTCGCG*, pero nuestra explicación fue más complicada y menos plausible que la que propuso Calladine.

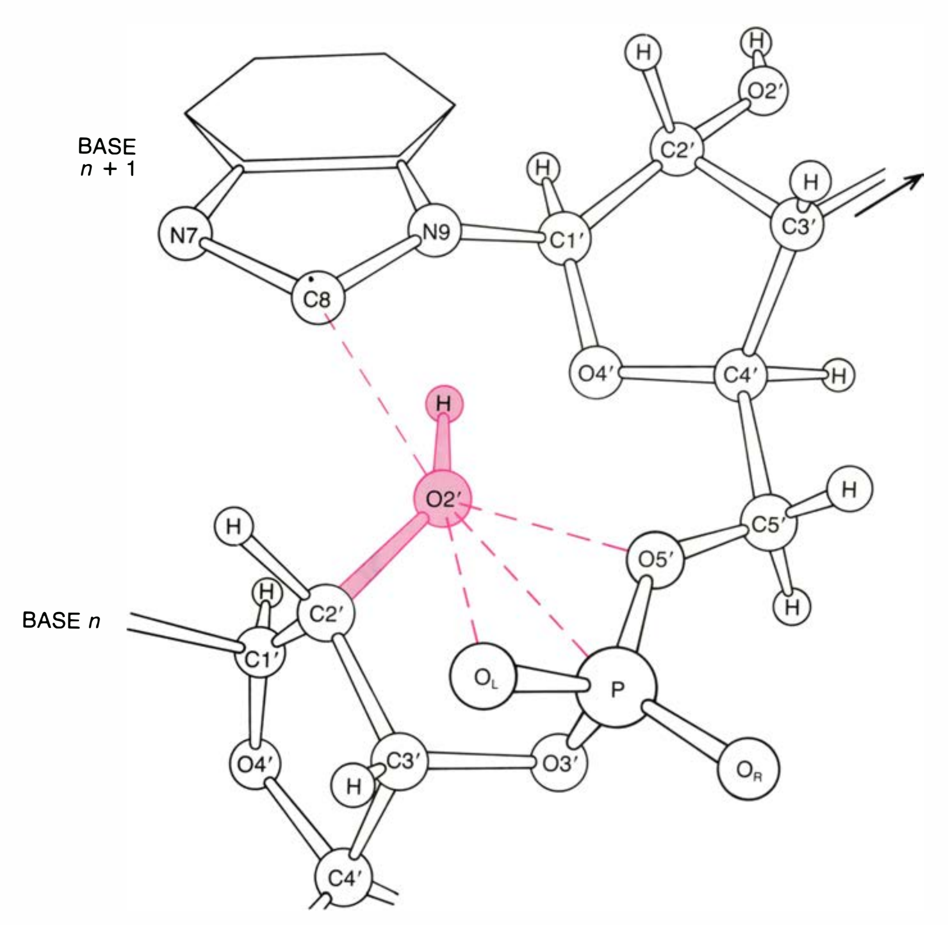
Predicción de la variación

Cualquiera que sea el borde del par de bases que choca con un vecino, siempre puede evitarse disminuyendo el enrollamiento local de la hélice entre esos dos pares de bases. Eso sugiere que la variación del enrollamiento puede tener la misma explicación estereoquímica sencilla que la del balanceo. El choque entre dos purinas de cadenas opuestas de la hélice puede evitarse directamente alabeando menos uno o los dos pares de bases, o, indirectamente, deslizando un par de bases a lo largo de su eje mayor, de manera que la purina se salga parcialmente del apilamiento helicoidal. Calladine concluyó de su estudio de este último efecto que los choques estéricos son aproximadamente el doble de pronunciados en el surco menor que en el mayor y requieren, por consiguiente, un reajuste doble de los parámetros locales de la hélice.

A partir de los supuestos de Calladine, nuestro grupo, que ahora trabaja en la Universidad de California en Los



13. PAR ESTEREOSCOPICO de una larga hélice *Z*, generada en el ordenador por reiteración de *GCGC*. Nótese que las estructuras de los ADN *A* y *Z* se oponen en muchos aspectos. La hélice *Z* es delgada y larga con un surco menor profundo, un surco mayor ancho y escaso alabeo. La hélice *A* es corta y rechoncha, con un surco mayor profundo, un surco menor superficial y pares de bases muy alabeados.

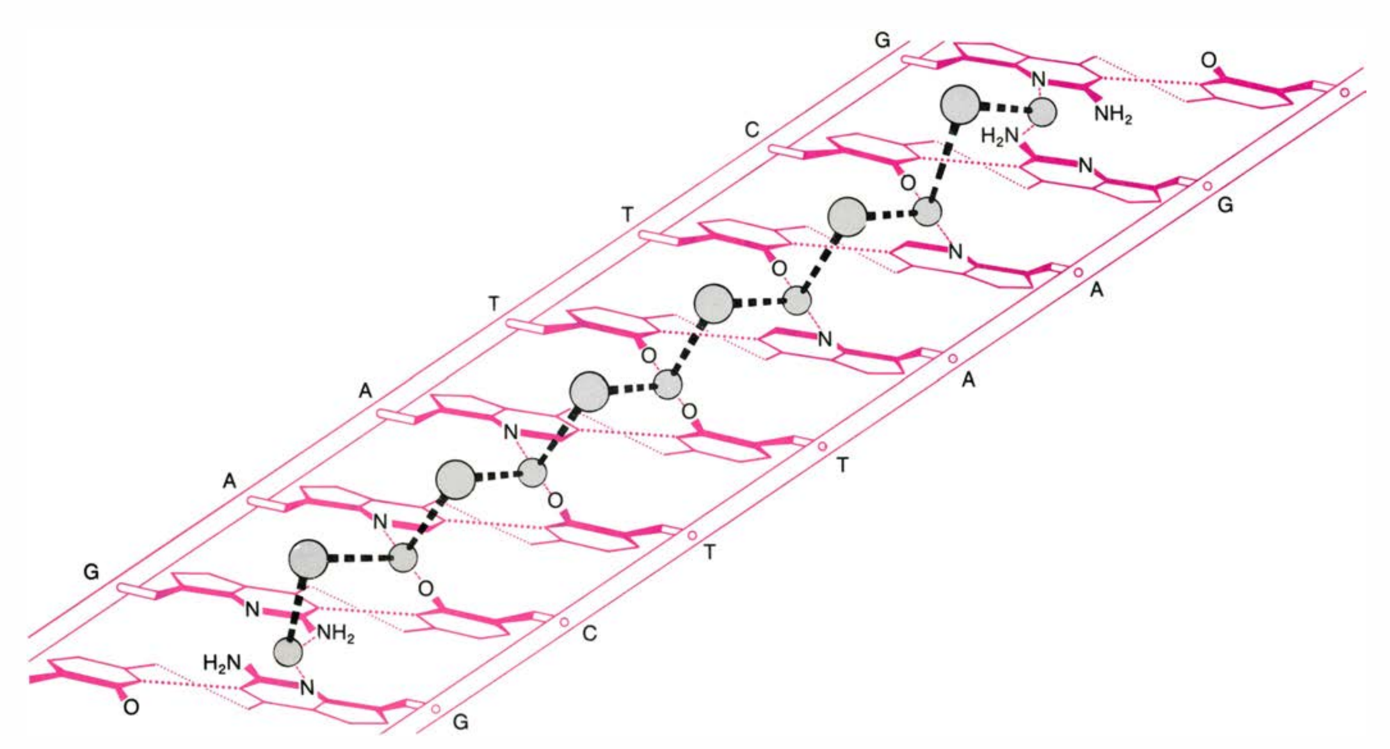


14. EL ARN no adopta la estructura *B* cuando forma una doble hélice autocomplementaria. La razón parece ser que la ribosa del ARN tiene un grupo hidroxilo (OH) donde la desoxirribosa del ADN sólo tiene un átomo de hidrógeno. Se ha añadido aquí un grupo OH (*color*) a una cadena de ADN *B* para simular un ARN *B*. Hay un contacto incómodo entre el átomo de oxígeno añadido y cuatro átomos del fosfato, el azúcar y la base adyacentes (*trazos*). En el ADN *A*, ese grupo 2′OH sobresale de la superficie de la hélice.

Angeles, ha desarrollado cuatro sencillas funciones sumatorias para predecir, a partir de la secuencia de bases del ADN, las variaciones locales de enrollamiento, balanceo, alabeo y también ángulo de torsión (que describe la conformación de la cadena alrededor del enlace C4′-C3′). Las funciones para el alabeo y la torsión sólo parecen aplicarse al ADN *B*, pero las dos primeras funciones valen para las hélices *A* y *B*. Como ejemplo de estas predicciones, la función Σ_1 para el enrollamiento se calcula como sigue:

1. A cada intervalo entre un par de bases y el siguiente se asigna un número que refleja la tendencia local relativa a reducir el enrollamiento para evitar choques estereoquímicos. Sea −2 este número para la secuencia purina-pirimidina que da lugar a choques en el surco menor, −4 para la secuencia pirimidina-purina, con su pronunciado choque en el surco menor, y cero para las secuencias purina-purina y pirimidina-pirimidina, que no tienen dificultades estéricas.

2. Se asignan números de magnitud la mitad y signo opuesto a los números anteriores a los intervalos contiguos, porque, al cerrar el espacio entre dos pares de bases girándolos alrededor de su eje mayor, cada uno de los intervalos contiguos se abrirá la mitad de lo que se cerró aquél. Obtenemos así, para la secuencia purina-pirimidina y los dos intervalos adyacentes los núme-



15. CORDON DE HIDRATACION (*negro*) del ADN *B*, formado por una masa zigzagueante de moléculas de agua. Aquí miramos hacia el surco menor de una hélice desenrollada del dodecámero de ADN *B*. Las cadenas están muy esquematizadas y, el alabeo, exagerado. La tanda inferior de moléculas de agua (*círculos grises pequeños*) se une por puentes de hidrógeno a los átomos de nitrógeno y oxígeno de los pares de bases adyacentes; otra tanda de moléculas de agua (*círculos grises mayores*) se une a la anterior para formar el cordón. El cordón impide que se abra el surco menor para formar la ancha depresión presente en el ADN *A* y se considera, por tanto, que es un factor importante para estabilizar la estructura *B* cuando el agua abunda.

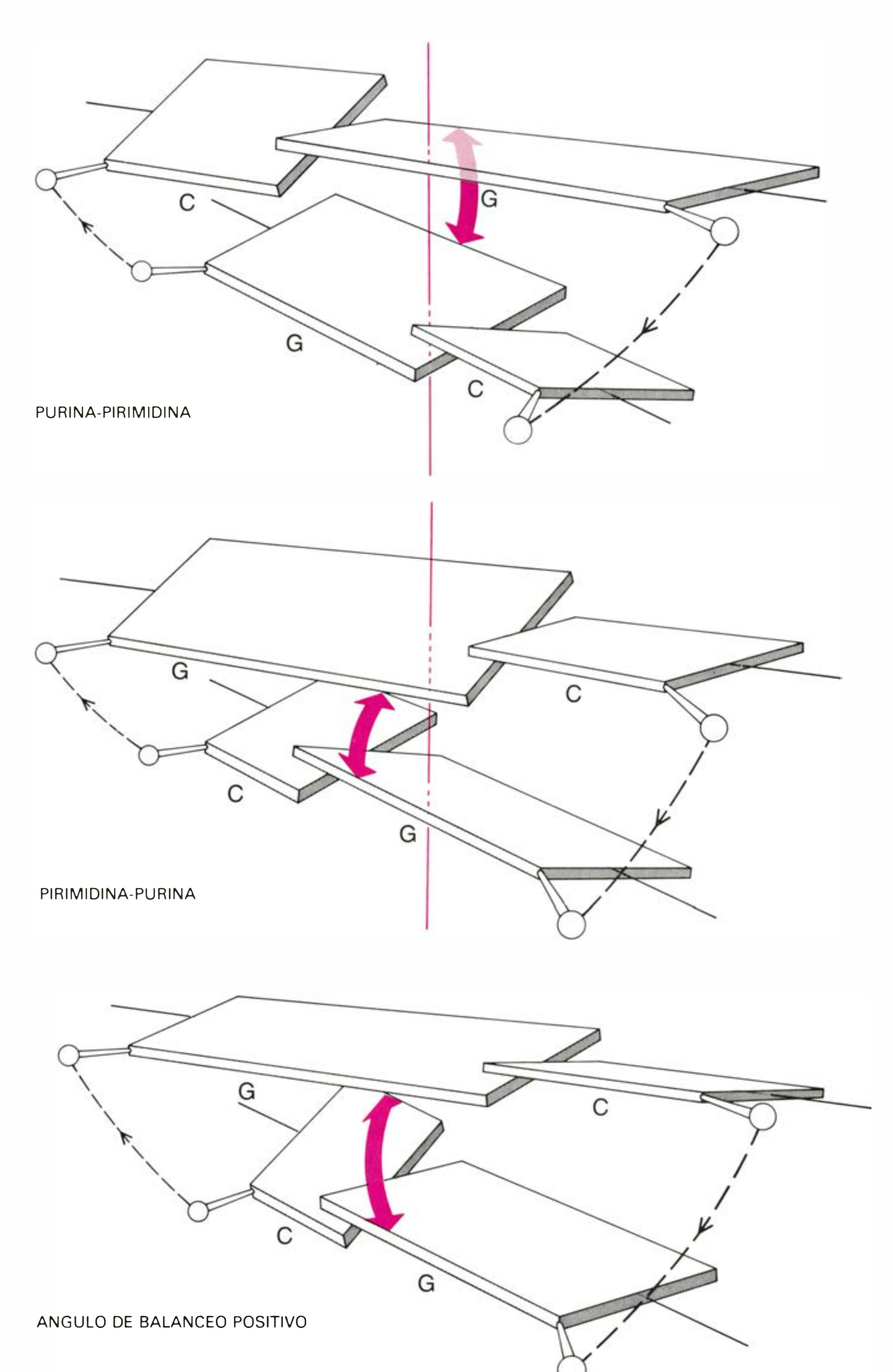
ros +1, −2, +1; para la secuencia pirimidina-purina, +2, −4, +2.

3. Se suman los valores asignados a cada intervalo para obtener la suma total Σ_1 que mide la variación local del enrollamiento.

La figura 17 explica la derivación del sumatorio Σ_1 para el dodecámero *CGCGAATTCGCG*, una hélice *B*, y compara los valores predichos y observados. La regresión lineal entre los ángulos observados y Σ_1 da un coeficiente de correlación de 0,994, o sea, una correlación muy significativa entre ambos. El sumatorio reproduce fielmente las variaciones de los ángulos observados, excepto en los extremos, en los que el solapamiento entre moléculas distintas del cristal distorsiona la hélice.

Se puede seguir un procedimiento parecido para calcular el sumatorio Σ_2, que mide la variación local del balanceo, excepto que las correcciones para la secuencia pirimidina-purina tienen unos signos opuestos (−2, +4, −2), porque la incompatibilidad estérica de esos pares de bases se evita separándolos por el lado del surco menor, es decir, haciendo el ángulo de balanceo más positivo. La derivación de Σ_2 y su comparación con los ángulos observados se ilustra en el caso de *CGCGAATTCGCG*. Se predice con precisión el balanceo de los pares de bases, excepto en los extremos de la hélice.

Si estas dos funciones sirvieran para una sola hélice de un solo tipo, el método sería como tantas otras ideas brillantes de la ciencia, que sólo funcionan una vez a manos del inventor. Resulta, sin embargo, que cuando se aplican las mismas reglas a las cuatro hélices *A* conocidas, los resultados son tan buenos como para la hélice *B*. Una de las estructuras *A* era una prueba particularmente dura. Los dos tetrámeros de *CCGG/CCGG* están apilados uno sobre otro en el cristal para formar una especie de octámero cuyos pares de bases centrales no estuvieran unidos por fosfatos. A pesar de esta falta de unión covalente, el enrollamiento y el balanceo de las bases en el intervalo central son exactamente los predichos por las funciones sumatorias para un octámero covalente intacto. Las fuerzas que apilan los pares de bases son, al parecer, más importantes que la cadena de azúcar y fosfato para determinar la estructura del ADN. La estructura *GGTATACC*, una hélice *A*, se ajusta también a las funciones del enrollamiento y el balanceo y sólo la función del enrollamiento empieza a fallar en el caso de los pares de bases híbridos de ADN y ARN de la hélice *(GCG)TATACGC*.



16. EL ALABEO de los pares de bases afecta a los contactos entre purinas de pares de bases adyacentes y cadenas opuestas. Las purinas y pirimidinas se representan con unas planchas, los átomos C1′ de los azúcares con pequeñas esferas y las cadenas están reducidas a líneas de trazos, con flechas que señalan la dirección 5′-3′. Si la secuencia es purina-pirimidina (*arriba*), el alabeo positivo acerca demasiado los bordes de las purinas por el lado del surco mayor (*flecha en color*). En la secuencia inversa (*centro*) el contacto inaceptable ocurre en el surco menor. Una posibilidad (*abajo*) de evitar el choque es balancear los pares de bases alrededor de sus ejes mayores, abriendo el ángulo que forman por el lado del choque.

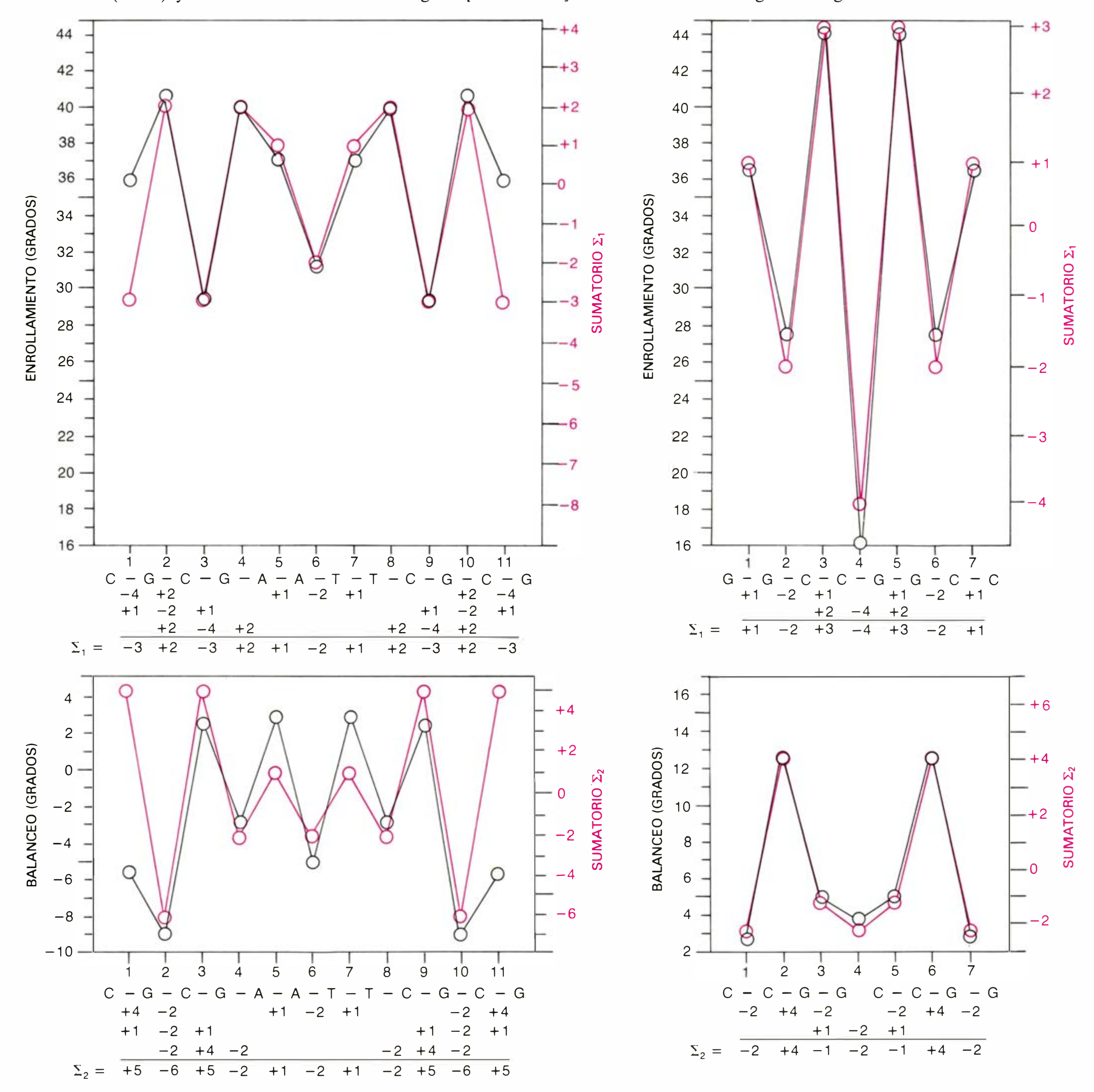
De la estructura a la regulación

¿Influyen las variaciones estructurales locales de la hélice en la lectura de la secuencia del ADN por represores y otras proteínas reguladoras? Si se acepta la idea de que la lectura implica la formación de puentes de hidrógeno entre las cadenas laterales de la proteína y los átomos de oxígeno y nitrógeno de los bordes de las bases, es difícil imaginar que variaciones de hasta 15 grados con respecto a la media no tengan un efecto importante en el proceso de reconocimiento. Quizá las fluctuaciones inducidas por la secuencia representan un ajuste fino del acoplamiento

entre proteína y ADN, que supera al mero reconocimiento de dadores y aceptores de puentes de hidrógeno en el borde de cada par de bases.

Tal extremo sólo se puede confirmar analizando con rayos X la estructura de un cristal de una proteína de reconocimiento unida a la secuencia de ADN correspondiente. Están en marcha varios estudios de tales complejos, que conciernen a los represores *lac, lambda* y *cro*, la proteína de la activación por catabolito (CAP) y varias restrictasas. John Rosenberg, de la Universidad de Pittsburgh, ha calculado recientemente un mapa de densidad electrónica de la restrictasa *Eco R*I unida a su lugar de reconocimiento en el dodecámero *CGCGAATTCGCG*. Ha encontrado que el ADN sigue estando en la forma *B*, pero distorsionado.

Lo que ya se sabe influye siempre en las preguntas que se pueden plantear. Un nuevo avance del conocimiento puede sacar un problema del reino de la vaga especulación y colocarlo en el centro de la investigación científica. Eso han hecho los análisis de rayos X de cristales con preguntas del siguiente tenor: ¿cómo afecta la secuencia de bases de un ADN bicatenario a la estructura y al comportamiento de la hélice? ¿Qué importancia tiene este efecto para la lectura de la información? Tardará en dárseles respuestas completas, pero vale la pena plantearse tales cuestiones, que nos llevarán a comprender mejor los procesos en que se apoya la regulación genética.



17. PREDICCION de las variaciones locales del enrollamiento y el balanceo a partir de la secuencia de bases con los sumatorios Σ_1 (enrollamiento) y Σ_2 (balanceo). Se aplica aquí el método a un dodecámero *B* (*dibujos de la izquierda*), un octámero *A* (*arriba a la derecha*) y un pseudooctámero *A* (*abajo a la derecha*). Las curvas negras y las escalas de la izquierda muestran los valores observados y, las curvas de color y las escalas de la derecha, las predicciones. A cada gráfica subyace el cálculo del sumatorio para cada intervalo entre pares de bases contiguos. Los valores observados y predichos coinciden muy bien: los coeficientes de correlación entre ellos son de 0,994 para el enrollamiento y 0,917 para el balanceo de la hélice *B*; 0,991 para el enrollamiento del octámero *A* y 0,995 para el balanceo del pseudooctámero *A*. El pseudooctámero (dos tetrámeros *A* apilados, no unidos por fosfatos) se comporta como un verdadero octámero; el apilamiento de los pares de bases parece más importante para la estructura local de la hélice que la continuidad de la cadena.

Gravedad cuántica

En una teoría mecánico-cuántica de la gravitación, la geometría real del espacio y el tiempo estaría sometida a continuas fluctuaciones, quedando incluso borrosa la distinción entre pasado y futuro

Bryce S. DeWitt

La gravedad parece ocupar un puesto especial entre las fuerzas de la naturaleza. Hay fuerzas, así el electromagnetismo, que actúan *dentro del* espacio-tiempo, mero marco para los sucesos físicos. Con la gravedad es distinto. No se trata de ninguna fuerza impuesta sobre un fondo espacio-temporal pasivo; antes bien, consiste en una distorsión del propio espacio-tiempo. Un campo gravitatorio es una "curvatura" del espacio-tiempo. Este es el concepto de gravedad al que llegó Einstein y que él mismo describía como el más difícil trabajo de su vida.

La distinción cualitativa entre la gravedad y las demás fuerzas resulta todavía más evidente cuando pretendemos formular una teoría de la gravitación que sea coherente con los preceptos de la mecánica cuántica. El mundo cuántico nunca está fijo. En la teoría cuántica del campo electromagnético, por ejemplo, siempre fluctúa el valor del campo. En un universo dominado por la gravedad cuántica, la curvatura del espacio-tiempo, e incluso la propia estructura, quedarían sujetas a fluctuaciones. Sin duda, cabe que el orden de los sucesos en el universo y el sentido de pasado y futuro sean susceptibles de cambio.

Podría argüirse que, si tales fenómenos existieran, seguramente se habrían observado ya. La realidad es que cualquier efecto explícito de la gravitación mecánico-cuántica queda confinado en unas dimensiones extremadamente pequeñas, dimensiones a las que prestó atención Max Planck en 1899. En ese año, Planck introdujo su famosa constante, llamada cuanto de acción y representada con la letra $\hbar$. Le preocupaba resolver el problema del espectro de radiación de un cuerpo negro (la luz que escapa a través de un pequeño agujero hecho en una cavidad caliente). Como nota curiosa al margen, observó que su constante, cuando se combinaba con la velocidad de la luz y la constante de la gravitación de Newton, establecía un sistema de unidades absoluto. Estas dimensiones fijaban la escala de la gravedad cuántica.

Las unidades de Planck no tienen nada que ver con el marco físico en que nos movemos diariamente. Recuérdese que su unidad de longitud es $1{,}61 \times 10^{-33}$ centímetros, o sea, 10^{21} veces más pequeña que el diámetro de un núcleo atómico. Con respecto a las dimensiones nucleares guarda más o menos la misma relación que la que hay entre el tamaño de un hombre y el de nuestra galaxia. La unidad de tiempo de Planck es todavía más fantástica: $5{,}36 \times 10^{-44}$ segundos. ¡Para trabajar experimentalmente con estas distancias y tiempos, usando instrumentos construidos con la tecnología actual, necesitaríamos un acelerador de partículas del tamaño de la galaxia!

Sin guía experimental que la oriente, la gravedad cuántica busca refugio en la especulación. Pero se mueve con un talante conservador. Toma las teorías existentes bien confirmadas, y se limita a considerarlas en sus conclusiones lógicas extremas. Reducida a su esqueleto esencial, la gravedad cuántica se propone combinar tres teorías: la relatividad restringida, la teoría de la gravitación de Einstein y la mecánica cuántica. Síntesis todavía por alcanzar en todos sus supuestos, si bien podemos aprender mucho de lo hasta ahora obtenido. Además, el desarrollo de una teoría operativa de la gravedad cuántica ofrece el único camino conocido para desentrañar el origen de la gran explosión (*big bang*) y el destino último de los agujeros negros, los sucesos que marcan el principio y el fin del universo.

De las tres teorías componentes de la gravedad cuántica, la primera por orden cronológico es la relatividad restringida. Une espacio y tiempo mediante el postulado (confirmado experimentalmente) que afirma que la velocidad de la luz es la misma para todos los observadores que se mueven en regiones del espacio vacías y libres de fuerzas exteriores. Para descubrir las consecuencias de este postulado, que fue enunciado por Einstein en 1905, se recurre a la ayuda de un diagrama de espacio-tiempo, gráfica cuyas curvas representan posiciones de los objetos en el espacio en función del tiempo. Dichas curvas reciben el nombre de líneas de universo.

Por mor de simplicidad, prescindiré de dos dimensiones del espacio. De ese modo, una línea de universo puede dibujarse en una gráfica bidimensional donde las distancias espaciales se miden horizontalmente y los intervalos de tiempo verticalmente. Una recta vertical es la línea de universo de un objeto que permanece en reposo en el sistema de referencia escogido para realizar las mediciones. Una recta inclinada es la línea de universo de un objeto que se mueve con velocidad constante en el sistema de referencia escogido. Una línea de universo curva corresponde a un objeto sometido a aceleraciones.

Un punto en el diagrama de espacio-tiempo nos da tanto una posición en el espacio como un instante de tiempo; recibe el nombre de suceso. La distancia espacial entre dos sucesos depende del sistema de referencia escogido; lo mismo le ocurre a un intervalo de tiempo. El concepto de simultaneidad depende, en realidad, del sistema de referencia. Dos sucesos que están en una misma línea horizontal, en cierto sistema de referencia, son simultáneos en dicho referencial, pero no lo son en otros referenciales.

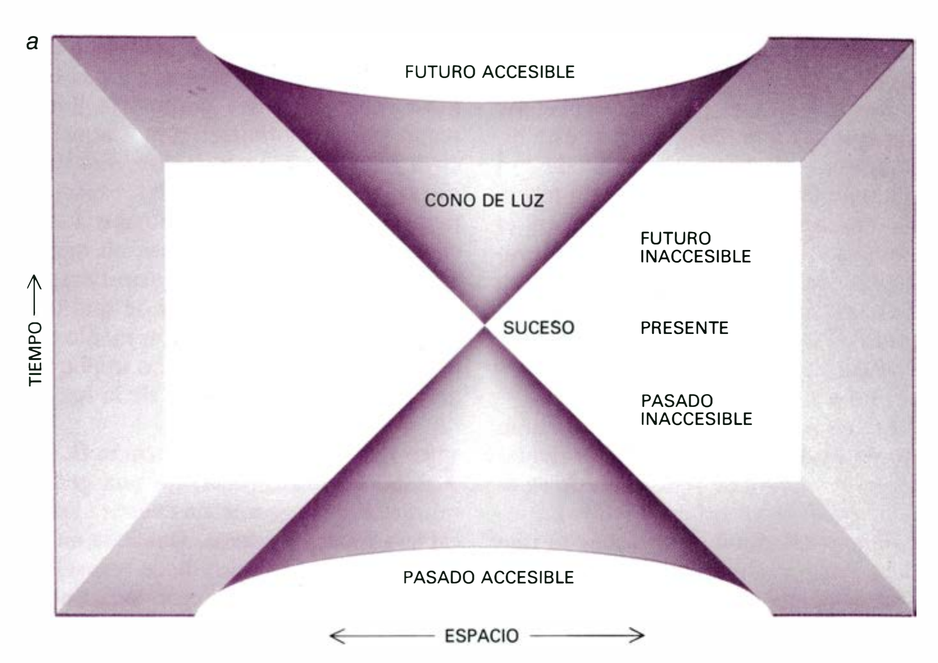
Para establecer una relación entre referenciales en movimiento relativo debemos introducir una unidad común para la medida del espacio y del tiempo. Como factor de conversión se utiliza la velocidad de la luz, relacionando

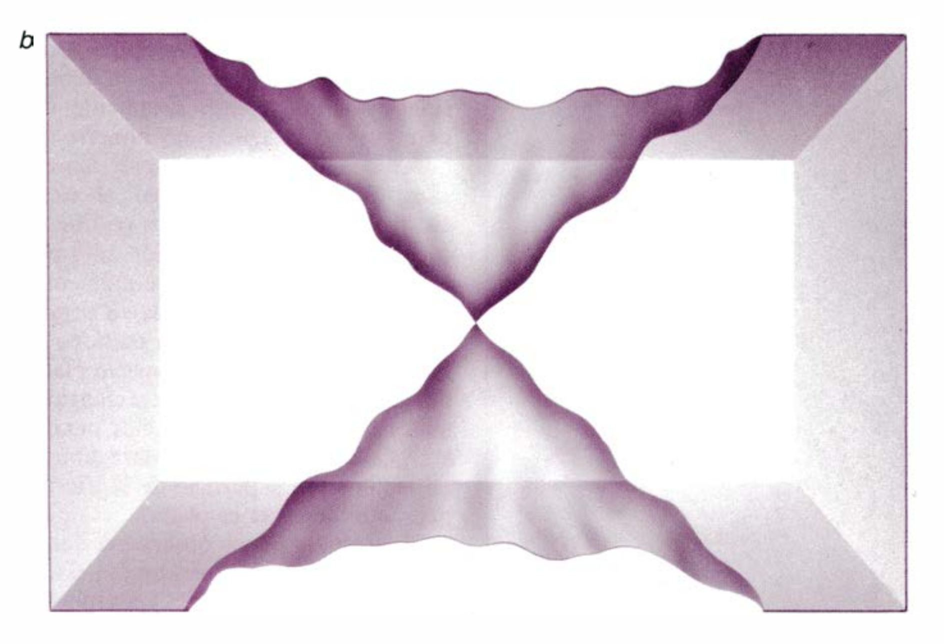
una distancia dada con el tiempo que la luz emplearía en recorrerla. A partir de aquí tomaré el metro como unidad, tanto para el espacio como para el tiempo. Un metro de tiempo equivale a unos 3,3 nanosegundos (milmillonésimas partes de un segundo).

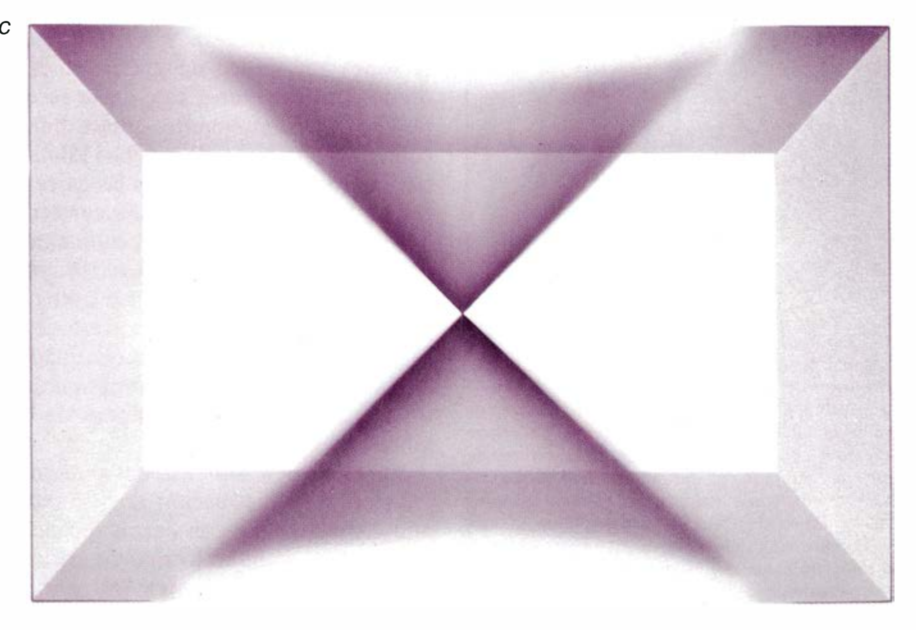
Cuando medimos el espacio y el tiempo con las mismas unidades, la línea de universo de un fotón (un cuanto de luz) tiene una pendiente de 45 grados. La línea de universo de cualquier objeto material debe mostrar una inclinación respecto de la vertical siempre menor que 45 grados, lo que es otra manera de decir que su velocidad es siempre menor que la de la luz. Si la línea de universo de cualquier objeto o señal se inclinara con ángulo superior a 45 grados respecto de la vertical, dicho objeto o señal parecería moverse hacia atrás en el tiempo para otros observadores. Estableciendo una serie de señales más rápidas que la luz, una persona podría transmitir información a su propio pasado, violando así el principio de causalidad. Dichas señales quedan proscritas en el cuadro de la relatividad restringida.

Consideremos dos sucesos en la línea de universo de un observador no acelerado. Supongamos que estos sucesos, en cierto sistema de referencia, están separados por cuatro metros de espacio y cinco metros de tiempo. Por tanto, en dicho referencial, el observador se mueve con una velocidad igual a cuatro quintas partes de la velocidad de la luz. En otro referencial, su velocidad sería diferente y también lo serían los intervalos de espacio y tiempo. Sin embargo, hay una cantidad que sería la misma en todos los sistemas de referencia. Esta cantidad invariante recibe el nombre de "tiempo propio" entre los dos sucesos; se trata del intervalo de tiempo que mediría un reloj que llevara el observador.

En el sistema de referencia escogido, la línea de universo entre los dos suce-



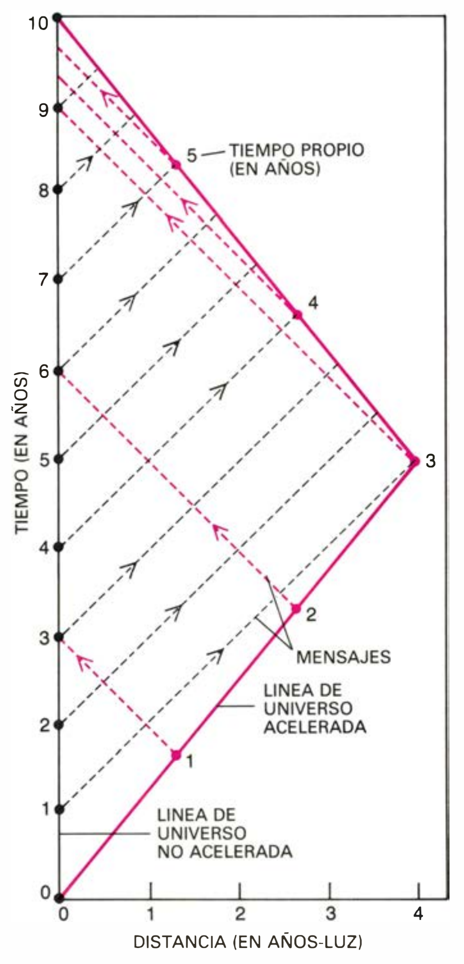




1. CONO DE LUZ, que define las regiones del universo accesibles desde un punto dado del espacio y desde un instante determinado del tiempo. Podría resultar un concepto mal definido en una teoría de la gravedad cuántica. El cono (*a*) es una superficie en el espacio-tiempo cuatridimensional, si bien aquí lo hemos dibujado suprimiendo una dimensión espacial. Si cuantificamos la gravitación, la forma del cono puede fluctuar desordenadamente para pequeñas distancias, (*b*). Las fluctuaciones, de hecho, no pueden percibirse directamente; el cono de luz aparece cual si estuviera cubierto de vello. Sólo podemos dar, pues, una respuesta en términos de probabilidad a la pregunta de si dos puntos del espacio-tiempo pueden comunicarse (por medio de señales que se muevan con velocidad no superior a la velocidad de la luz) (*c*).

sos es la hipotenusa de un triángulo rectángulo cuya base mide cuatro metros y, cinco, su altura. El "tiempo propio" corresponde a la "longitud" de esta hipotenusa, aunque calculada de un modo no usual: por medio de un teorema "pseudopitagórico". Como en el teorema de Pitágoras, primero elevamos al cuadrado los catetos. Ahora bien, en relatividad restringida, el cuadrado de la hipotenusa es igual a la diferencia de los cuadrados de los catetos, no a su suma.

En el presente ejemplo el tiempo propio es de tres metros. Sigue siendo así en el sistema de referencia de cualquier observador no acelerado. La invariancia del tiempo propio es lo que une espacio y tiempo en una entidad, el espacio-tiempo. La geometría del espacio-tiempo, basada en un teorema pseudo-pitagórico, no es euclídea, pero en muchos aspectos es análoga. En geometría euclídea, de entre todas las líneas que unen dos puntos, la línea recta puede definirse como aquella cuya longitud es un extremo. Esto es cierto en la geometría del espacio-tiempo. Sin embargo, en geometría euclídea el extremo es siempre un mínimo, mientras que en el espacio-tiempo es un máximo, a condición de que los dos puntos puedan unirse por medio de una línea de universo que no implique velocidades superiores a la de la luz.



2. LINEA DE UNIVERSO, que define un camino a través del espacio y el tiempo. En el dibujo hay dos líneas de universo que ilustran una versión de la paradoja de los mellizos de Einstein. La línea de universo "doblada", que corresponde al mellizo que sufre una aceleración en el punto de retorno de su viaje, parece la más larga; la verdad es que este mellizo es el que emplea un "tiempo propio" menor. En un diagrama de espacio-tiempo, la línea recta representa siempre el máximo intervalo entre dos puntos. Señalados están los tiempos de emisión y llegada de los mensajes emitidos por los mellizos.

En 1854, el matemático alemán G. F. B. Riemann introdujo una generalización de la geometría euclídea aplicable a espacios curvos. Desde la antigüedad se habían estudiado espacios curvos en dos dimensiones. Se llaman superficies curvas y se abordan desde la perspectiva de un espacio euclídeo ordinario de tres dimensiones. Riemann demostró que un espacio curvo podía tener un número cualquiera de dimensiones y estudiarse intrínsecamente. No es necesario imaginárselo inmerso en un espacio euclídeo de mayor dimensión. Riemann señaló que el espacio físico en el que vivimos podría ser curvo. Para él, esta cuestión sólo podría contestarse experimentalmente. ¿Cómo acometer, al menos en principio, su ensayo? El espacio euclídeo es un espacio plano. En un espacio plano puede dibujarse una malla rectangular uniforme por medio de rectas paralelas. ¿Qué pasaría si se intentara dibujar tal malla en la superficie de la Tierra, pensando que ésta es plana?

El resultado puede observarse desde un avión que sobrevuele, un día claro, las regiones cultivadas de las Grandes Praderas (Great Plains) en Estados Unidos. Las tierras están subdivididas, por medio de carreteras que van en las direcciones este-oeste y norte-sur, en regiones que miden un kilómetro cuadrado. Las carreteras este-oeste se extienden frecuentemente como líneas continuas a lo largo de muchos kilómetros, pero no sucede así con las carreteras norte-sur. Siguiendo una carretera hacia el norte se observan, cada cierto número de kilómetros, saltos abruptos hacia el este o el oeste. Estos saltos se producen a causa de la curvatura de la Tierra. Si los saltos fueran eliminados las carreteras se acercarían mutuamente, creando secciones de menos de un kilómetro cuadrado.

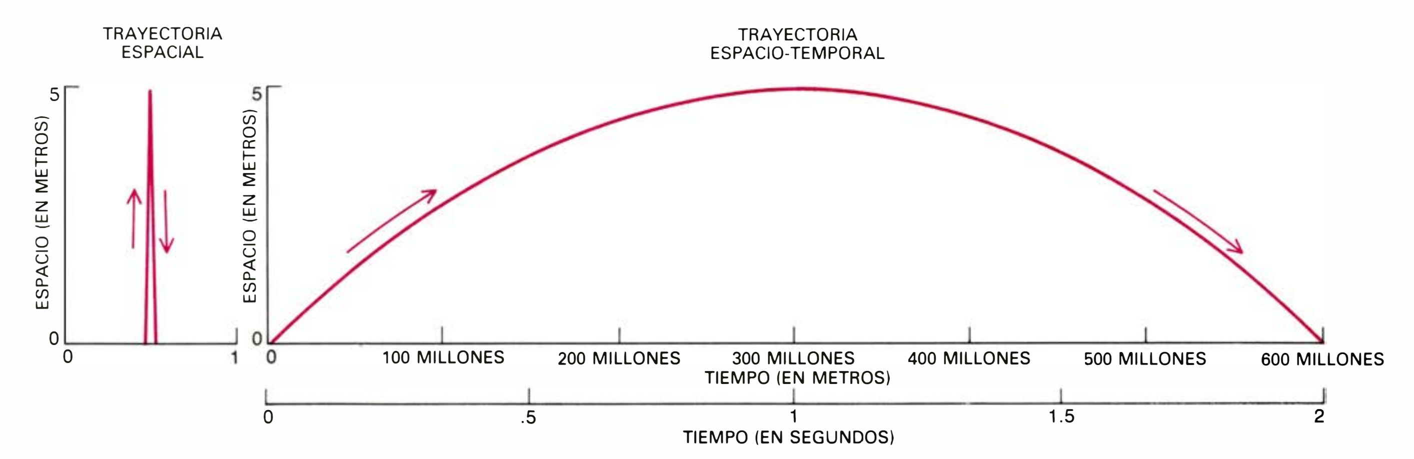
En el caso tridimensional podemos imaginar la construcción de un andamio gigante hecho con barras rectas de igual longitud unidas formando ángulos exactamente iguales a 90 y 180 grados. Con un espacio plano, la construcción del andamio no presentaría ninguna dificultad. Pero si es curvo tendríamos que acabar por acortar las barras o alargarlas para que encajaran.

La misma generalización que Riemann introdujo en la geometría euclídea puede aplicarse a la geometría de la relatividad restringida. Dicha generalización fue realizada por Einstein, entre 1912 y 1915, con la ayuda del matemático Marcel H. Grossmann. El resultado es una teoría del espacio-tiempo curvo. Einstein la desarrolló pensando en una teoría de la gravitación. En la relatividad restringida se supone que no hay campos gravitatorios y que el espacio-tiempo es plano. En un espacio-tiempo curvo existe un campo gravitatorio; en realidad "curvatura" y "campo gravitatorio" son sinónimos.

Por ser su teoría gravitatoria una generalización de la relatividad restringida, Einstein la llamó relatividad general. Denominación que induce a error. La verdad es que la relatividad general es menos relativista que la relatividad restringida. La ausencia completa de estructura en el espacio-tiempo plano, su homogeneidad e isotropía nos aseguran que las posiciones y velocidades sean estrictamente relativas. En cuanto el espacio-tiempo presente "protuberancias", regiones locales con curvatura, será absoluto, puesto que posición y velocidad podrán especificarse con respecto a estas protuberancias. El espacio-tiempo no se limita a ofrecer un escenario sin estructura para la física; está dotado él mismo de propiedades físicas.

En la teoría de Einstein la materia produce la curvatura. La relación que hay entre cantidad de materia y valor de la curvatura, simple en teoría, exige un cálculo complicado. Para describir la curvatura en un punto se necesitan veinte funciones de las coordenadas de este punto. De ellas, diez corresponden a una parte de la curvatura que se propaga libremente en forma de ondas gravitatorias u "ondulaciones de la curvatura". Las otras diez funciones quedan determinadas por la distribución de la masa, energía, momento, momento angular y tensiones internas de la materia, así como por la constante gravitatoria de Newton, G.

Comparada con las densidades másicas que se dan en la Tierra, G es una constante muy pequeña. Se necesita una gran cantidad de masa para curvar el espacio-tiempo de modo apreciable. El recíproco, $1/G$, puede utilizarse como una medida de la "rigidez" del



3. CURVATURA DEL ESPACIO-TIEMPO, debida a la existencia de una masa; es lo que caracteriza al campo gravitatorio. (En un espacio-tiempo curvo existe un campo gravitatorio; "curvatura" y "campo gravitatorio" son sinónimos.) Cuando lanzamos una pelota a una altura de cinco metros (*izquierda*), permanece arriba durante dos segundos. Su movimiento de subida y bajada pone de manifiesto la curvatura del espacio-tiempo cerca de la superficie de la Tierra. Aunque la curvatura de la trayectoria se aprecia con toda nitidez, la verdad es que resulta muy pequeña cuando medimos espacio y tiempo en las mismas unidades. Por ejemplo, podemos convertir los segundos en metros si los multiplicamos por la velocidad de la luz, o sea, 300 millones de metros por segundo. La trayectoria nos daría un arco de 600 millones de metros de largo y cinco de flecha (*derecha*). La altura aparece exagerada.

espacio-tiempo. Dentro de las condiciones experimentales habituales, el espacio-tiempo resulta muy rígido. Toda la masa de la Tierra produce una curvatura del espacio-tiempo que es sólo de una milmillónesima parte de la curvatura de la superficie de la Tierra.

En la teoría de Einstein un cuerpo que cae u orbita libremente sigue una línea de universo geodésica. Una geodésica que une dos puntos del espacio-tiempo es una línea de universo cuya longitud presenta un extremo entre ellos; constituye una generalización de la noción de línea recta. Si imaginamos el espacio-tiempo curvo inmerso en un espacio plano de mayor dimensión, esa geodésica constituirá una línea curva.

El efecto de la curvatura sobre un cuerpo que se mueve se ha comparado, con frecuencia, al movimiento de una bola que rueda sobre una lámina deformada de material elástico. El modelo puede prestarse a confusiones, ya que sólo representa curvaturas espaciales. En la vida real estamos obligados a permanecer en un universo de cuatro dimensiones con espacio y tiempo. No podemos evitar movernos en este universo, pues estamos lanzados hacia adelante en el tiempo. El tiempo es el elemento clave. Resulta que, aún siendo curvo el espacio en un campo gravitatorio, la curvatura del tiempo es mucho más importante. La razón estriba en el valor grande de la velocidad de la luz, que es la magnitud que relaciona la escala de espacio con la de tiempo.

En la vecindad de la Tierra la curvatura del espacio, ligerísima, no se detecta con medidas estáticas. Lo que no empece que nuestra precipitada carrera en el tiempo sea tan rápida que, en situaciones dinámicas, la curvatura resulte apreciable, como un pequeño bache en una autopista, que, inadvertido por el peatón, se convierte en un peligro para un automóvil a gran velocidad. Aunque el espacio cerca de la Tierra parece plano incluso con un alto grado de precisión, podemos ver la curvatura del espacio-*tiempo* simplemente lanzando una pelota al aire. Si la pelota permanece dos segundos en el aire sigue un arco de una altura de cinco metros. La luz, en dos segundos, recorre 600.000 kilómetros. Si imaginamos el arco de cinco metros alargado horizontalmente hasta los 600.000 kilómetros, la curvatura de este arco da la curvatura del espacio-tiempo.

La introducción, por Riemann, de la idea de los espacios curvos abrió el estudio de otra fecunda área matemática: la topología. Se sabía que las superficies sin borde de dimensión dos se presentaban en infinitos tipos diferentes que no podían deformarse continuamente de unos a otros; dos ejemplos sencillos son la esfera y el toro. Riemann señaló que lo mismo es cierto para espacios curvos de mayor dimensión y dio los primeros pasos para clasificarlos.

Existe también una diversidad infinita de tipos topológicos de espacio-tiempo curvos (o, más precisamente, de modelos de espacio-tiempo curvos). A la hora de describir un universo real, se descartarán varios modelos, por conducir a paradojas sobre la causalidad o por no admitir leyes físicas conocidas. Pero queda todavía un sorprendente número de posibilidades.

Alexander A. Friedmann, matemático ruso, propuso en 1922 un modelo interesante de universo. En relatividad restringida, el espacio-tiempo se toma no sólo plano sino también infinito en extensión, tanto para el espacio como para el tiempo. En el modelo de Friedmann, cada sección espacial tridimensional del espacio-tiempo tiene volumen finito y topología de una 3-esfera, espacio que puede ser inmerso en un espacio euclídeo de cuatro dimensiones, de tal modo que todos sus puntos equidistan de un punto dado. Desde el descubrimiento, en los años veinte, por Edwin P. Hubble, de la expansión del universo, este modelo ha gozado del favor de los cosmólogos. Cuando combinamos el modelo de Friedmann con la teoría de la gravitación de Einstein, se puede predecir una gran explosión en un instante inicial de compresión infinita, seguida de una expansión que va frenándose a lo largo de miles de millones de años, debido a la mutua atracción gravitatoria existente entre toda la materia del universo.

Un espacio-tiempo de Friedmann tiene la propiedad de que toda curva cerrada, contenida en él, puede reducirse de modo continuo a un punto. Al espacio-tiempo dotado de esta propiedad se le llama conexo. El universo real puede carecer de esa propiedad. El modelo de Friedmann parece describir muy bien la región del espacio que está a unos pocos miles de millones de años-luz de nuestra galaxia, pero no podemos ver entero al universo.

Un universo cuya estructura se repitiera indefinidamente, como un dibujo en un papel de pared, en una dirección espacial dada, sería un buen ejemplo de un universo múltiplemente conexo. En un universo así cada galaxia es un miembro de una serie infinita de galaxias idénticas separadas por una distan-

cia fija (y necesariamente enorme). Si los miembros de una serie son exactamente idénticos, es cuestionable el que se consideren distintos. Más rentable resulta pensar que cada serie representa una sola galaxia. De este modo, el viaje desde un miembro de la serie al próximo devuelve al viajero a su punto de partida, y la línea trazada por tal viaje es una línea curva cerrada e irreductible a un punto. Es como una curva cerrada en la superficie de un cilindro rodeándolo una vez. El universo que se repite recibe el nombre de universo cilíndrico.

El "agusanamiento", introducido en 1957 por John Archibald Wheeler, hoy en la Universidad de Texas en Austin, constituye otro ejemplo, a una escala francamente menor, de una estructura múltiplemente conexa. Puede construirse un agusanamiento de dos dimensiones cortando dos agujeros circulares en una superficie bidimensional y uniendo continuamente los bordes de los cortes [*véase la figura 5*]. El procedimiento es el mismo en tres dimensiones, sólo que más difícil de visualizar.

Los dos agujeros pueden estar muy distanciados en el espacio original y, no obstante, muy juntos a través de la "garganta" que los conecta; el agusanamiento se ha popularizado en ciencia-ficción en un ingenio que permite ir de un sitio a otro mucho más rápido que lo haría la luz: basta pinchar dos agujeros en el espacio, conectarlos y recorrer la garganta. Por desgracia, aunque lográramos construir un punzón adecuado (lo que es dudoso), el esquema no funcionaría. Si la geometría del espacio-tiempo se rige por las ecuaciones de Einstein, el agusanamiento es un objeto dinámico. Los dos agujeros que une son, necesariamente, agujeros negros y cuanto penetra en ellos ya no puede salir. Ocurre que la garganta se estrecha: todo lo que contiene se comprime hasta densidades infinitas antes de que pueda llegar al otro lado.

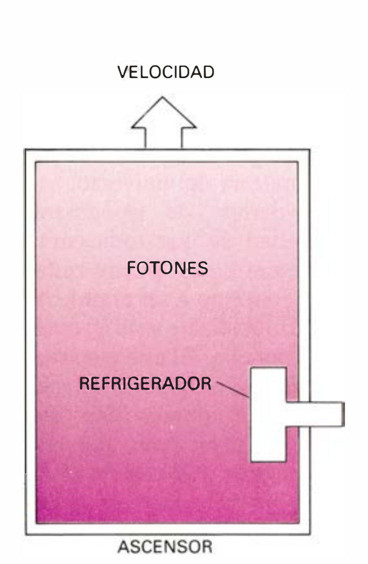
La tercera componente de la gravedad cuántica, la mecánica cuántica, fue introducida en 1925 por Werner Heisenberg y Erwin Schrödinger. Mas, en su formulación inicial, no tomaba en consideración la relatividad restringida. Lo que restó brillo a su espectacular éxito desde el mismo comienzo; esperaban explicación multitud de observaciones experimentales en las que los efectos cuánticos predominan y en las que la relatividad desempeña un papel menor o despreciable. Con todo, se sabía que los electrones de algunos átomos alcanzan velocidades que dan una fracción apreciable de la velocidad de la luz; de aquí que no tardara mucho en llegar una teoría cuántica relativista.

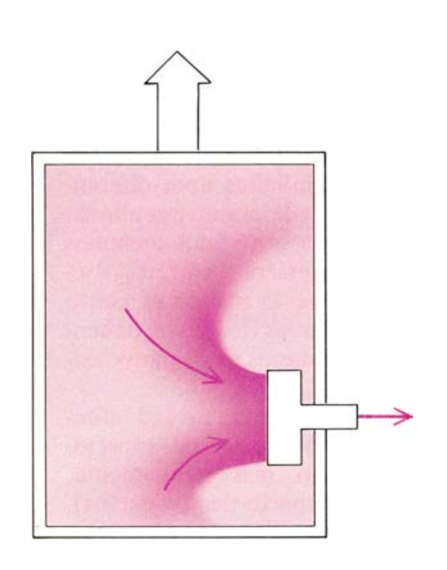
Hacia la mitad de los años treinta se sabía ya que, cuando se combina la teoría cuántica con la relatividad, se deduce una serie de hechos completamente nuevos. Dos son los fundamentales. Primero, cada partícula queda asociada con un tipo de campo y cada campo se asocia con una clase de partículas indistinguibles. Los campos electromagnético y gravitatorio no podían considerarse ya los únicos campos fundamentales de la naturaleza. Segundo, hay dos tipos de partículas, clasificadas según su momento angular de espín (cuantificado). Aquellas cuyo espín vale $1/2\ \hbar$, $3/2\ \hbar$, y así sucesivamente, obedecen al principio de exclusión (no hay dos en el mismo estado cuántico); aquellas cuyo espín vale 0, $\hbar$, $2\hbar$, y así sucesivamente, son gregarias.

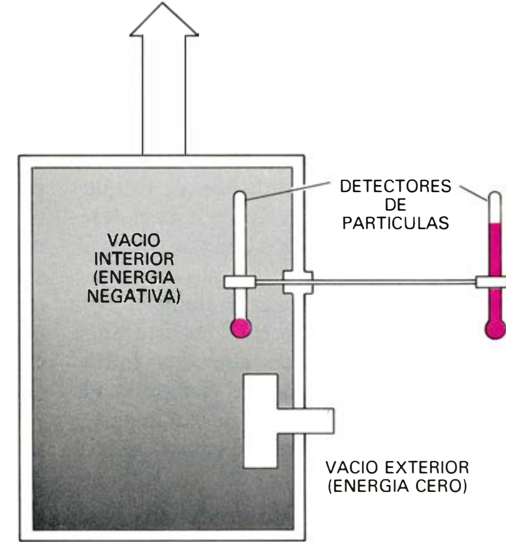
Estas asombrosas consecuencias derivadas de la reunión de la relatividad restringida y la mecánica cuántica han sido repetidamente confirmadas en el último medio siglo. La relatividad y la teoría cuántica dan juntas una teoría que es mayor que la suma de sus partes. Este efecto sinergético se ve aún más aumentado cuando añadimos la gravedad.

En física clásica plana, al espacio-tiempo desocupado se le llama vacío. El vacío clásico carece de estructura. En física cuántica, se da el nombre de vacío a una entidad mucho más compleja, con una rica estructura. Esta estructura surge de la existencia en el vacío de campos libres no nulos, es decir, campos alejados de sus fuentes.

Un campo electromagnético libre es matemáticamente equivalente a una colección infinita de osciladores armónicos, que pueden representarse como masas unidas a muelles. En el vacío, cada oscilador se halla en su estado fun-







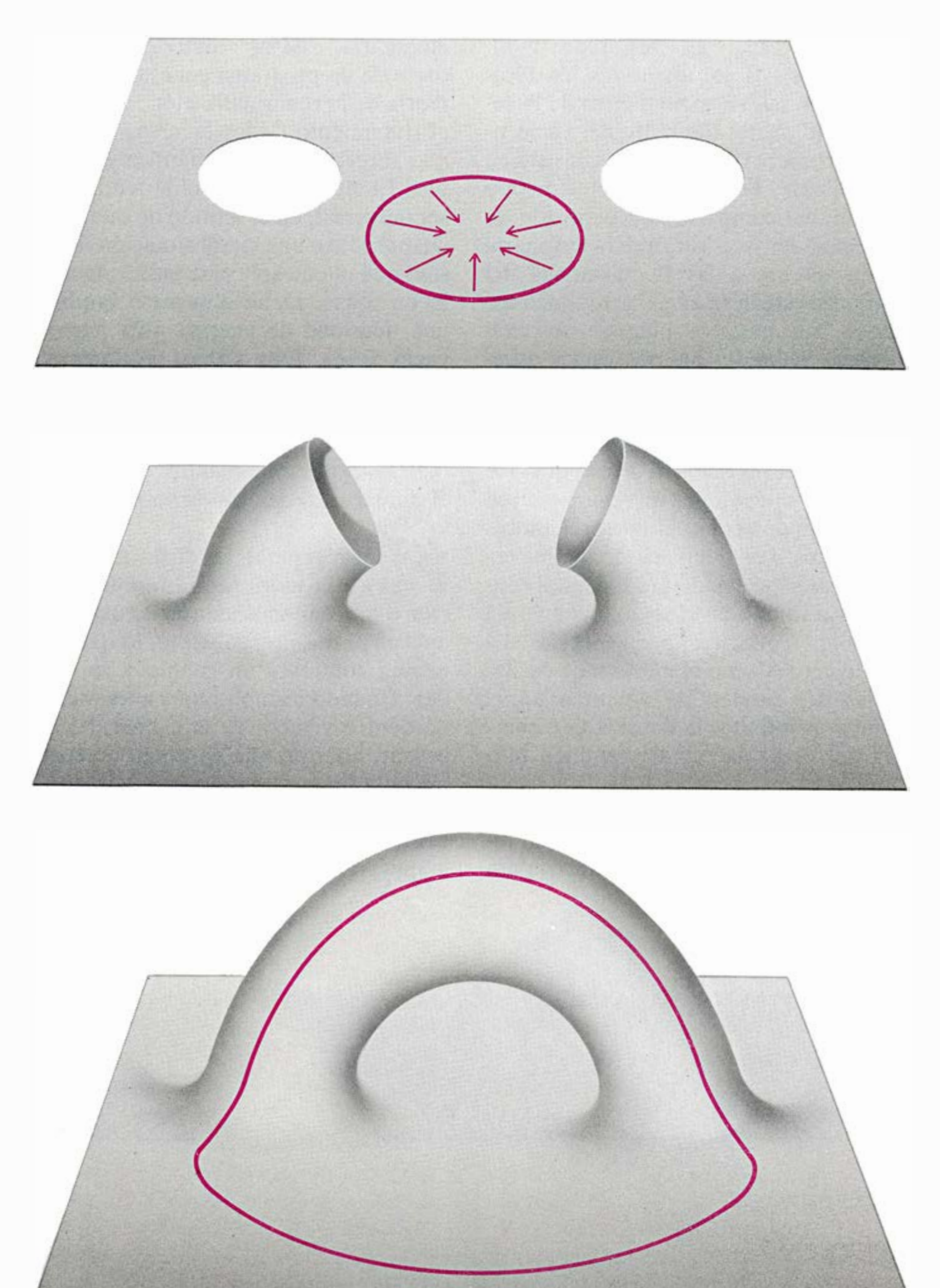
4. CABINA DE ASCENSOR ACELERADA, que sirve para plantear un experimento ideal que pone de manifiesto la naturaleza del vacío en mecánica cuántica y el efecto que producen en el vacío tanto la aceleración como la gravedad. Suponemos que la cabina, sin nada, es hermética; al principio, el vacío es perfecto, fuera y dentro de la cabina. En cuanto se inicia la aceleración, el suelo emite una onda electromagnética y la cabina se llena de un gas diluido de fotones, cuantos de radiación electromagnética (*izquierda*). Un refrigerador que utiliza una fuente de energía externa saca los fotones (*centro*). Eliminados los fotones, dos detectores de éstos miden la energía del vacío dentro y fuera de la cabina (*derecha*). Como el instrumento del exterior está acelerado en el vacío, es sensible a las fluctuaciones mecánico-cuánticas de los campos que impregnan el espacio incluso en ausencia de partículas. El detector interior está en reposo respecto a la cabina; no detecta fluctuaciones. De aquí deducimos que los vacíos dentro y fuera de la cabina no son equivalentes. Si al vacío "habitual", que es el exterior, le asignamos por definición una energía cero, el vacío del interior poseerá una energía negativa. Deberíamos devolver los fotones eliminados por el refrigerador para traer la energía a cero. Un campo gravitatorio puede crear un vacío con energía negativa.

damental, el de más baja energía. Cuando un oscilador clásico (no mecánico-cuántico) se encuentra en su estado fundamental, permanece quieto en una posición bien definida. Esto no es cierto para un oscilador cuántico. Si un oscilador cuántico estuviera en un punto concreto, se conocería su posición con infinita precisión; entonces, el principio de incertidumbre implicaría que tiene energía y momento infinitos, lo que es imposible. En el estado fundamental de un oscilador cuántico no están fijados con precisión ni la posición ni el momento. Ambos valores permanecen sujetos a fluctuaciones aleatorias. En el vacío cuántico es el campo electromagnético (y cualquier otro campo) el que fluctúa.

Las fluctuaciones del campo en el vacío cuántico, aleatorias, son de un tipo especial. Satisfacen el principio de relatividad, de modo que "parecen" las mismas a cualquier observador no acelerado, independientemente de su velocidad. Puede demostrarse que esta propiedad implica que el campo es cero en promedio y que la amplitud de las fluctuaciones aumenta cuando disminuye la longitud de onda. El resultado final es que un observador no puede determinar su velocidad por medio de las fluctuaciones.

Sin embargo, las fluctuaciones pueden utilizarse para determinar la aceleración. En 1976, William G. Unurh, de la Universidad de British Columbia, demostró que un hipotético detector de partículas, sometido a una aceleración constante, reaccionaría frente a las fluctuaciones del vacío como si estuviera quieto dentro de un gas de partículas (y por tanto, no en el vacío) que tuviera una temperatura proporcional a la aceleración. Un detector sin acelerar no reacciona frente a las fluctuaciones.

La idea de que la temperatura y la aceleración pueden relacionarse de este modo condujo a un replanteamiento de lo que se entiende por "vacío" y a reconocer que existen varias clases de vacíos. Uno de los vacíos atípicos más sencillos puede crearse repitiendo, dentro de un contexto mecánico-cuántico, un experimento ideal propuesto ya por Einstein. Se trata de un ascensor cerrado que circula libremente por el espacio vacío. Un "espíritu juguetón" empieza a tirar de él hasta que adquiere una aceleración constante en la dirección suelo-techo. Se supone que las paredes de la cabina son conductores perfectos, impermeables a la radiación electromagnética; la cabina se supone completamente vacía, sin ninguna partícula. Einstein introdujo este escenario imaginario como un modo de ilustrar la equivalencia de la gravitación y la aceleración; reconsiderándolo, nos daremos cuenta de que pueden esperarse también varios efectos típicamente mecánico-cuánticos.



5. AGUSANAMIENTO EN EL ESPACIO-TIEMPO, una hipotética entidad que podría cambiar la topología del universo. En un plano, produciremos un agusanamiento cortando dos agujeros y estirando los bordes para formar tubos que luego unimos. En el plano inicial, toda curva cerrada podía ser reducida a un punto (*color*). No podemos hacerlo con una curva que atraviese el agusanamiento. Un agusanamiento en un espacio tridimensional o en un espacio-tiempo cuatridimensional no son conceptualmente distintos.

Para empezar, en el momento en que se inicia la aceleración el suelo de la cabina emite una onda electromagnética que se propaga hasta el techo y luego oscila de uno a otro. (La razón de la emisión de la onda se evidencia en un análisis matemático detallado de un conductor eléctrico acelerado, pero el efecto es análogo al de creación de una onda de compresión acústica, que aparecería con la cabina llena de aire.) Si permitimos, temporalmente, cierta disipación en las paredes de la cabina, la onda electromagnética se convierte en fotones con un espectro energético térmico, o en otras palabras, en la radiación característica de un cuerpo negro a cierta temperatura.

La cabina contiene ahora un gas diluido de fotones. Podemos instalar un refrigerador, con un radiador fuera, para eliminar los fotones, a costa de una cierta energía suministrada por una

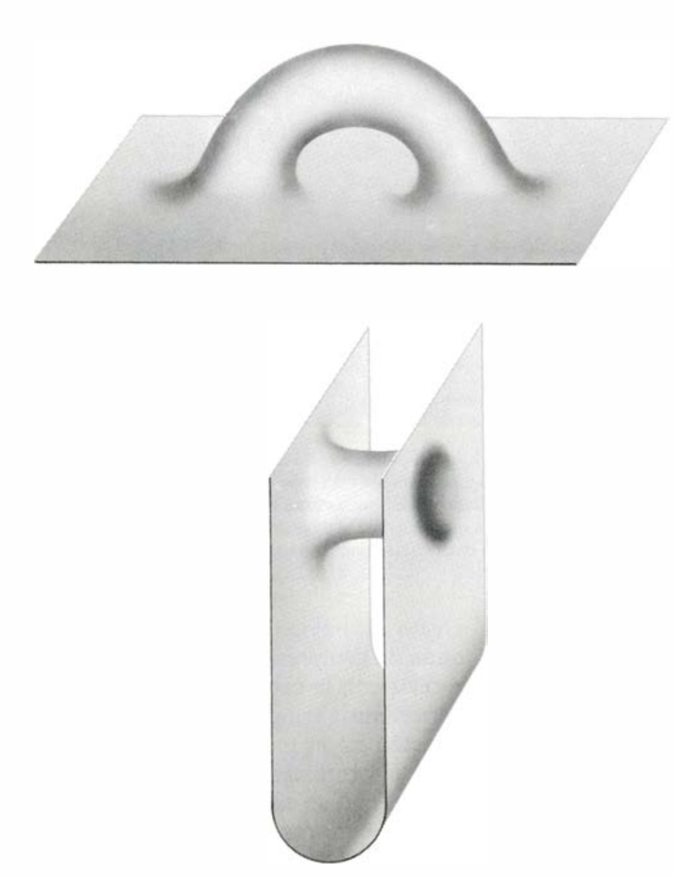
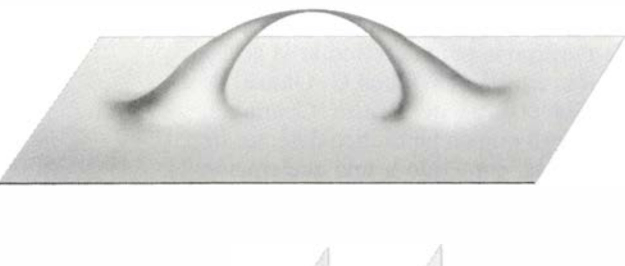
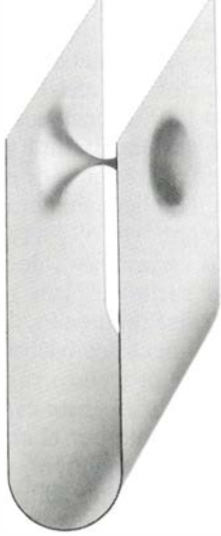
fuente externa. El resultado final, cuando hemos sacado todos los fotones, es un nuevo vacío dentro de la cabina, sutilmente diferente del vacío típico exterior. La diferencia consiste en lo siguiente. Primero, un detector tipo Unruh que comparta la aceleración de la cabina del ascensor, que reaccionaría térmicamente a las fluctuaciones del campo si estuviera en el vacío típico exterior, no muestra ninguna relación dentro. Segundo, los dos vacíos difieren en su contenido energético.

Especificar la energía del vacío obliga a resolver algunos problemas delicados de la teoría cuántica de campos. He señalado antes que un campo libre equivale a una colección de osciladores armónicos. Las fluctuaciones de los osciladores en el estado fundamental confieren al campo en el vacío una energía residual, conocida por energía del punto cero. Como el número de osciladores por unidad de volumen en el campo es infinito, la densidad de energía del vacío parece ser también infinita.

Una densidad de energía infinita es un estorbo. Los físicos teóricos han introducido varios métodos técnicos para eliminarla. Estos métodos forman parte de un programa general, llamado teoría de la renormalización, destinado al tratamiento de los diversos infinitos que surgen en teoría cuántica de campos. El método que se elija tendrá que ser universal, en el sentido de que no se adapte sólo a una cierta situación física, sino que pueda aplicarse uniformemente en todas. Debe obtenerse también una densidad de energía nula para el vacío típico. Esta última condición es esencial para que se muestre coherente con la teoría de Einstein, ya que el vacío típico es el equivalente cuántico del espacio-tiempo plano y vacío. Si en él hubiera energía, no sería plano.

Por regla general, las distintas versiones de la teoría de la renormalización dan resultados idénticos cuando se aplican al mismo problema, lo que nos inspira confianza en cuanto a su validez. Cuando los aplicamos a los vacíos de dentro y fuera de la cabina del ascensor, aportan una densidad de energía nula fuera y una densidad de energía negativa dentro. Sorprende una densidad de energía negativa. ¿Qué es lo que puede ser menos que nada? Una breve reflexión nos dará el sentido de este valor negativo. Debemos añadir fotones térmicos al interior de la cabina para conseguir que el detector tipo Unruh del interior se comporte como si estuviera en el vacío típico exterior. Cuando añadimos estos fotones, su energía hace subir la total interior hasta cero, es decir, la hace igual a la del vacío exterior.

Hay que señalar que será difícil observar en la práctica estos extraños efectos. Para las aceleraciones comunes en la vida corriente, incluso en máquinas que alcanzan grandes velocidades, la energía negativa resulta demasiado pequeña para ser detectada. Sin embargo, hay un caso en el que se ha detectado la energía negativa del vacío, por lo menos indirectamente: en un efecto predicho, en 1948, por H. B. G. Casimir, de los Laboratorios Philips, en los Países Bajos. En el efecto Casimir, se colocan muy juntas en el vacío dos placas metálicas planas, microscópicas, descargadas, paralelas y limpias. Se atraen débilmente con una fuerza que puede atribuirse a la densidad de energía negativa del vacío que hay entre ellas.

6. REGIONES ALEJADAS DEL UNIVERSO podrían unirse, en principio, por medio de agusanamientos; ello nos sugiere que podrían establecerse, entre las mismas, comunicaciones más rápidas que la luz; la verdad es que tal esquema no funciona. En el dibujo del primer agusanamiento (*arriba a la izquierda*), la distancia entre los agujeros desde el "mundo exterior" es comparable a la que hay a través de la "garganta". En el tercer agusanamiento (*abajo a la izquierda*), la distancia exterior es mucho mayor. En los dibujos inferiores, el espacio representado por el plano parece curvado, pero ello obedece a que se ve desde una perspectiva de un espacio de mayor dimensión; al observador que viviera en el plano le parecería llano. Carece de sentido preguntarse si la garganta es o no un atajo; no se puede pasar por ella. La razón estriba en que un agusanamiento conecta siempre dos agujeros negros. La garganta "se estrecha" (*a la derecha*) y todo cuanto entra en la misma queda comprimido hasta una densidad infinita antes de alcanzar el otro lado.

Con un espacio-tiempo curvo, el vacío resulta todavía más complejo. La curvatura influye sobre la distribución espacial de las fluctuaciones del campo cuántico y, como la aceleración, puede originar una energía no nula en el vacío. Puesto que la curvatura puede variar de un punto a otro, la energía del vacío lo hará también, apareciendo negativa en unos sitios y positiva en otros.

En cualquier teoría coherente, la energía debe conservarse. Supongamos, por un momento, que un aumento en la curvatura cause un aumento en la energía del vacío cuántico. Este aumento debe provenir de alguna parte y, por tanto, la misma existencia de fluctuaciones en el campo cuántico implica que necesitemos energía para curvar el espacio-tiempo. Deducimos, pues, que el espacio-tiempo se opone a curvarse. Tal ocurre en la teoría de Einstein.

En 1967, el físico ruso Andrei Sakharov propuso que la gravitación podría ser un fenómeno puramente cuántico, proviniente de la energía del vacío, y sugirió que la constante de Newton G, o lo que es lo mismo, la rigidez del espacio-tiempo, podría calcularse a partir de unos primeros principios. Esta propuesta se enfrenta a varias dificultades. Primera, requiere que la gravedad sea reemplazada, como campo fundamental, por algún "campo de aforo de gran unificación", sugerido por las partículas elementales conocidas. Debería introducirse una masa fundamental para obtener, aquí también, una escala absoluta de unidades; así pues, una constante fundamental quedaría reemplazada por otra.

Segunda, y tal vez más importante, la dependencia respecto de la curvatura que resulta para la energía del vacío conduce a una teoría de la gravedad más complicada que la de Einstein. Dependiendo del número y tipo de campos elementales que se escojan y del método de renormalización, la energía del vacío, en vez de crecer cuando aumenta la curvatura, puede incluso decrecer. Esta relación inversa implicaría que el espacio-tiempo es inestable y que tendería a arrugarse como una ciruela. Aquí, supondré que el campo gravitatorio es fundamental.

Un vacío genuino constituye un estado de equilibrio térmico a la temperatura del cero absoluto. En la gravedad cuántica, dicho vacío sólo puede existir si la curvatura no depende del tiempo. Cuando la curvatura depende del tiempo, pueden aparecer espontáneamente partículas en el vacío (con el resultado, por supuesto, de dejar de ser un vacío).

El mecanismo a través del cual se producen las partículas puede explicarse nuevamente en términos de osciladores armónicos. Cuando la curvatura del espacio-tiempo cambia, las propiedades físicas de los osciladores del campo cambian también. Supongamos que uno de los osciladores está, al principio, en su estado fundamental, sometido a oscilaciones del punto cero. Si cambiamos alguna de sus propiedades, como la masa o la rigidez del muelle, también las oscilaciones alrededor del punto cero deben ajustarse al cambio. Después de realizado el ajuste, existe una probabilidad finita de que el oscilador no se halle ya en el estado fundamental, sino en un estado excitado. Este fenómeno es análogo al incremento en la vibración inducido en la cuerda oscilante de un piano cuando aumentamos la tensión. Dicho efecto recibe el nombre de excitación paramétrica. En el campo cuántico, el análogo de la excitación paramétrica es la producción de partículas.

Las partículas generadas por la variación con el tiempo de la curvatura aparecen aleatoriamente. No es posible predecir exactamente dónde o cuándo se va a crear una determinada partícula. Lo que sí podemos calcular es la distribución estadística de la energía y el momento de las partículas. La producción de partículas es máxima allí donde la curvatura es máxima y donde cambia más rápidamente. Seguramente fue muy grande durante la gran explosión y podría haber influido sustancialmente en la dinámica del universo en sus momentos iniciales. ¡Podría ser que las partículas creadas en estas circunstancias explicaran la aparición de toda la materia del universo!

Hace unos diez años que comenzaron los estudios sobre cálculo de producción de partículas en la gran explosión. Los iniciaron el académico ruso Yakov B. Zel'dovich y Leonard E. Parker, de la Universidad de Wisconsin en Milwaukee, cada uno por su cuenta. Desde entonces muchos otros científicos lo han intentado. Aunque algunos de los resultados son alentadores, ninguno es definitivo. Además, hay una grave cuestión, de la que todo depende: ¿cuál es el estado cuántico inicial que debemos escoger en el instante de la gran explosión? El físico debe ponerse entonces en el lugar del Dios creador. Ninguna de las propuestas hechas hasta ahora parece totalmente convincente.

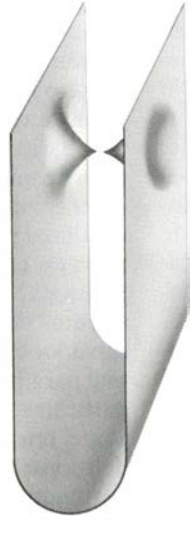
El otro suceso del universo en el transcurso del cual la curvatura varía rápidamente es el colapso de una estrella, formando un agujero negro. En este caso, los cálculos mecánico-cuánticos han dado un resultado sorprendente, prácticamente independiente de las condiciones iniciales. En 1974, Stephen W. Hawking, de la Universidad de Cambridge, demostró que la variación de la curvatura cerca de un agujero negro que colapsa crea un flujo radiante de partículas. Este flujo es estacionario y persiste largo tiempo después de que el agujero negro haya quedado geométricamente inactivo. Puede continuar porque el tiempo parece dilatarse en la zona de enorme intensidad del campo gravitatorio cercana al "horizonte" del agujero negro; para un observador exterior parece haber una virtual paralización de toda actividad. Las partículas creadas muy cerca del horizonte retrasan notablemente el inicio de su viaje hacia afuera.

Aunque el retraso en la emisión indica la existencia de una enorme cantidad de partículas acumuladas cerca del horizonte, "esperando que les llegue su turno" para escapar, la densidad de energía total en esta región es, en realidad, negativa y más bien pequeña. La energía positiva que corresponde a las partículas queda sobradamente compensada por la enorme energía negativa del vacío que existiría si no hubiera partículas (por ejemplo, si el agujero negro hubiera existido siempre y no lo hubiera creado un colapso gravitatorio).

Puede demostrarse que las partículas emitidas son estadísticamente independientes y que su espectro de energía es térmico. El carácter de cuerpo negro que tiene la radiación es, tal vez, su propiedad más importante. Permite asignar a cada agujero negro una temperatura y una entropía. La entropía, que mide el desorden termodinámico del sistema, resulta proporcional al área del horizonte. Para un agujero negro con masa estelar es enorme, mayor que 10^{19} multiplicado por el valor que tenía la entropía de la estrella que ha colapsado para formar el agujero negro. Por otro lado, la temperatura es inversamente proporcional a la masa y, con masa estelar, hay que dividir la temperatura de la estrella de origen por un número mayor que 10^{11}.

Como la cantidad de radiación emitida por un objeto depende de su temperatura, la radiación de Hawking procedente de un agujero negro astrofísico es totalmente despreciable. Resulta importante sólo para los "mini" agujeros negros, aquellos que tienen una masa inferior a 10^{10} gramos. El único modo





7. **TOPOLOGIA FLUCTUANTE, una característica del espacio-tiempo de la que sacan partido algunas formulaciones de la gravedad cuántica; presenta grandes dificultades conceptuales. Aquí se muestran dos casos de un agusanamiento que acaba de romperse, dando dos "hoyuelos". Si este suceso puede acaecer, el proceso inverso debe ser también posible; es decir, los hoyuelos deberían lograr unirse y formar un nuevo agusanamiento. Este proceso inverso parece plausible cuando los hoyuelos están cerca uno de otro, pero no cuando están alejados. Sin embargo, las ideas de "cerca" y "lejos" dependen de cómo se realiza la inmersión de la superficie en el espacio de mayor dimensión. Para un observador dentro de la superficie, los dos casos representados en el dibujo serían indistinguibles. (Ilustración de George V. Kelvin.)**

imaginable por el que pueden haberse creado miniagujeros negros es mediante la compresión que se produjo durante la gran explosión. En ese momento cabe que se se produjeran en gran cantidad; en cuyo caso habrían contribuido de un modo importante al valor de la entropía del universo.

La energía de las partículas creadas por el cambio con el tiempo de la curvatura no puede aparecer de la nada. Se toma del propio espacio-tiempo. De ello se deduce que las partículas actúan sobre el espacio-tiempo. Se ha intentado calcular esta "retroacción" en el caso de la gran explosión, así como determinar su efecto dinámico sobre el universo primitivo. Entre los objetivos marcados, uno ha sido el de averiguar si esta retroacción puede suprimir la densidad inicial infinita de materia requerida por la teoría clásica de Einstein. La densidad infinita actúa de barrera para toda pregunta posterior. Si puede reemplazarse simplemente por una enorme densidad, podríamos inquirir: ¿en qué consistía el universo *antes* de la gran explosión?

En la década de los sesenta, Roger Penrose, de la Universidad de Oxford, y el propio Hawking demostraron que la teoría clásica de Einstein era incompleta. Esta teoría predice la existencia, hacia el pasado o el futuro, de densidades infinitas y de curvaturas infinitas a partir de una extensa gama de condiciones físicas actuales razonables. Una teoría que predice un valor infinito para una cantidad observable deja de ser apta para predecir algo más allá de ese punto. Como los físicos pretenden la comprensión total de la naturaleza, esperan que una teoría así pueda ser mejorada, a fin de incluir una más amplia variedad de fenómenos. Desde un punto de vista cauto se afirma que la inclusión de los efectos cuánticos es la única solución razonable para evitar la no completitud de la teoría de Einstein.

Los cálculos de la retroacción durante la gran explosión se hacen por medio de simulaciones numéricas con un ordenador digital. Hasta ahora, los resultados obtenidos son ambiguos. Entre las dificultades a resolver estaba la de introducir, como dato para el ordenador, un valor razonable para la densidad de energía combinada de las partículas producidas y del vacío cuántico en el que quedan superpuestas.

El efecto de la retroacción adquiere especial importancia en el caso de un agujero negro. La radiación de Hawking disminuye tanto la energía como la entropía del agujero negro. Por consiguiente, la masa del agujero disminuye. Al principio, la velocidad de disminución es lenta, pero sube con la temperatura. Al final, la velocidad es tan grande que las aproximaciones hechas en los cálulos de Hawking no resultan aceptables. No se sabe qué sucede a partir de ese punto. Hawking piensa que sus aproximaciones mantienen validez cualitativa y que el final de la vida del agujero negro consiste en un destello espectacular, dejando momentáneamente como residuo una "singularidad desnuda" en la estructura causal del espacio-tiempo.

Cualquier singularidad, desnuda o no, representa un fracaso de la teoría. Si Hawking tiene razón, no sólo la teoría de Einstein es incompleta, sino que también lo es la teoría cuántica. La razón estriba en que, por cada partícula que se produce fuera del horizonte, otra se crea dentro. Las dos partículas están correlacionadas, en el sentido de que un observador podría detectar "efectos de interferencia en sus probabilidades" si pudiese comunicarse simultáneamente con ambas. Hawking supone que las partículas interiores son aplastadas hasta densidad infinita y dejan de existir. En ese momento falla la típica interpretación estadística de la mecánica cuántica. El propio significado de la probabilidad se pierde también en este aplastamiento infinito.

De acuerdo con una hipótesis alternativa, igualmente plausible, el propio marco de la teoría cuántica de campos que se construye alrededor de la teoría de Einstein evita que se pierdan, en el colapso, la probabilidad y la información. Es muy posible que la retroacción sea tan fuerte que evite un aplastamiento infinito. El horizonte, que es más una construcción matemática que física, puede que no esté allí cual barrera estrictamente unidireccional. Podemos saber qué pasa, partícula por partícula, con la masa que colapsa para dar el agujero negro. Nadie duda de que debe haber al final un destello de radiación de Hawking y una enorme densidad dentro del agujero. Con todo, las fuertes presiones a que están sometidas las partículas nucleares pueden convertirlas en fotones u otras partículas sin masa, que terminarían por escapar, llevándose la poca energía restante y todas las correlaciones cuánticas. No es preciso que estos productos finales se lleven nada de la entropía del agujero negro, que habría sido enteramente agotada por la radiación de Hawking.

Llego ahora a la parte más difícil y oscura de la gravedad cuántica. Cuando un efecto cuántico, como la producción de partículas o la energía del vacío, retroacciona sobre la curva-

tura del espacio-tiempo, la propia curvatura se transforma en un objeto cuántico. Un marco teórico coherente exige la cuantificación del propio campo gravitatorio. Para longitudes de onda largas comparadas con la longitud de Planck, las fluctuaciones cuánticas del campo gravitatorio cuantificado son pequeñas. Podemos representarlas adecuadamente si las tratamos como una débil perturbación del fondo clásico. La perturbación puede analizarse del mismo modo que lo haríamos con un campo independiente. Dicha perturbación contribuye a la energía del vacío y a la producción de partículas.

Para longitudes de onda y energías del orden de las de Planck la situación es mucho más complicada. Las partículas asociadas a un campo gravitatorio débil se llaman gravitones; carecen de masa y su momento angular de espín vale $2\hbar$. No es probable que podamos ver nunca los gravitones directa e individualmente. La materia ordinaria, incluso toda una galaxia de materia ordinaria, es casi totalmente transparente para ellos. Sólo cuando alcanzan energías de Planck interaccionan de modo apreciable con la materia. Sin embargo, a estas energías, son capaces de inducir curvaturas del orden de Planck en la geometría de fondo. El campo con el que están asociadas ya no es débil y el propio concepto de "partícula" está mal definido.

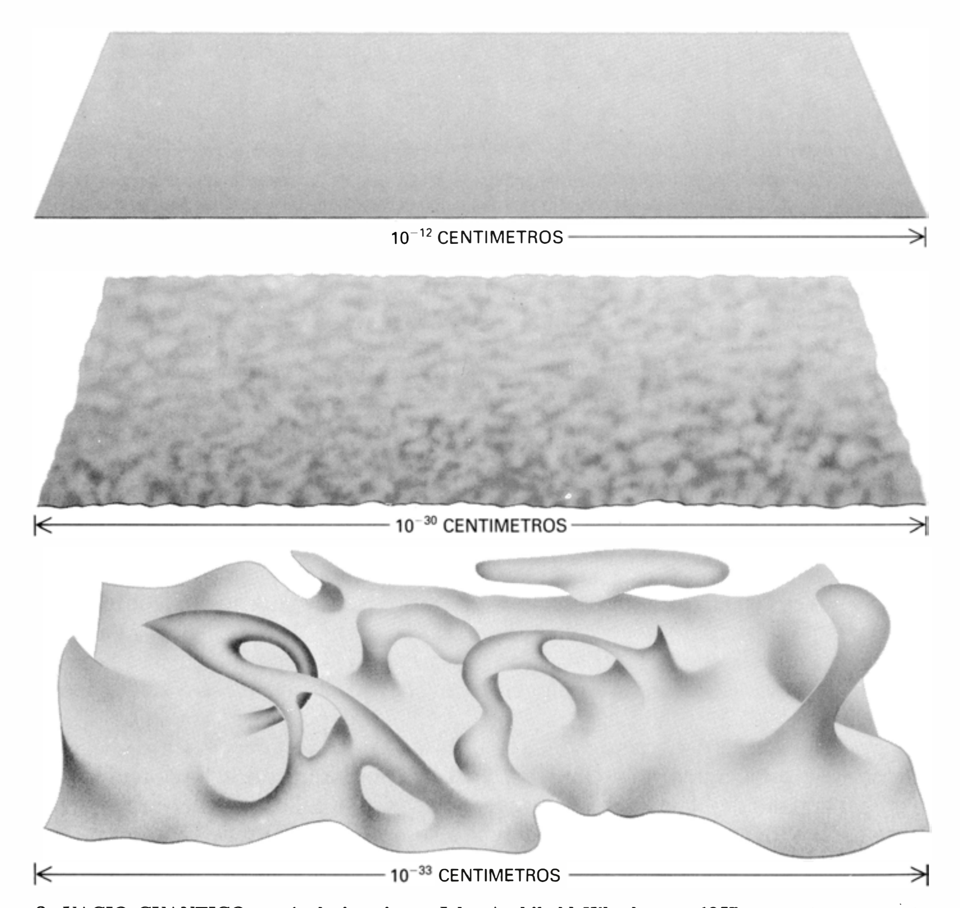
Para longitudes de onda largas, la energía que lleva el gravitón distorsiona la geometría de fondo; para longitudes de onda cortas, distorsiona incluso las ondas asociadas con el gravitón. Esto es debido a la no linealidad de la teoría de Einstein: cuando superponemos dos campos gravitatorios, el campo resultante no es igual a la suma de las dos componentes. Todas las teorías de campo que no sean triviales son no-lineales. En algunas se puede tratar la no-linealidad por medio de aproximaciones sucesivas, lo que se llama teoría de perturbaciones, nombre que se ha tomado prestado a la mecánica celeste. La clave del método está en mejorar una aproximación inicial mediante una serie de correcciones cada vez más pequeñas. Cuando aplicamos la teoría de perturbaciones a campos cuánticos, nos encontramos con infinitos que deben ser eliminados por renormalización.

En el caso de la gravedad cuántica la teoría de perturbaciones no funciona. Existen dos motivos. Primero, a las energías de Planck, los términos sucesivos de la serie de perturbaciones (o sea, las correcciones sucesivas) son comparables en magnitud. Cortar la serie cuando tenemos un número finito de términos no nos da una aproximación válida; lo que hay que hacer es sumar los infinitos términos de la serie. Segundo, los términos de la serie no pueden ser renormalizados individualmente de un modo coherente. En cada orden de aproximación aparecen nuevas clases de infinitos, que no se corresponden con ninguno de los de la teoría cuántica de campos ordinaria. Su aparición se debe a que, cuando cuantificamos el campo gravitatorio, cuantificamos también el espacio-tiempo. En la teoría cuántica de campos ordinaria el espacio-tiempo es un fondo fijo. En la gravedad cuántica el fondo no sólo se ve afectado por las fluctuaciones cuánticas sino que él mismo las tiene.

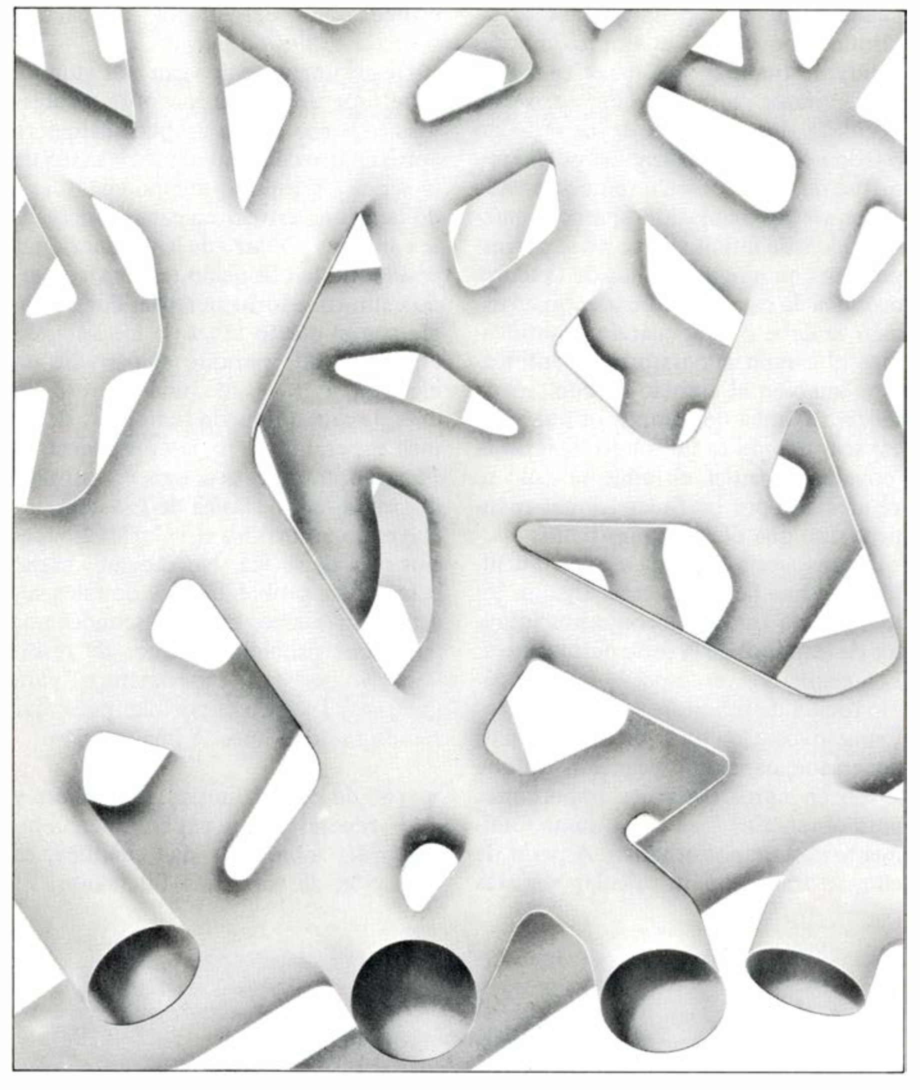
En respuesta limitada a estas dificultades se han llevado a cabo algunos intentos para sumar subconjuntos infinitos de términos de la serie perturbativa. Los resultados, en especial la total desaparición de infinitos, son alentadores y sugestivos. Con todo, estos resultados deben tomarse con precaución, ya que se hacen aproximaciones importantes para obtenerlos y nunca se suma totalmente la serie perturbativa. A pesar de ello, se utilizan para calcular mejores estimaciones del efecto retroactivo sobre la explosión inicial.

Desde un punto de vista más amplio, cabe esperar que surjan otros problemas cuya resolución no pueda siquiera aproximarse con la ayuda de la suma de la serie. Un espacio-tiempo cuantificado tiene una estructura causal fluctuante e incierta. Dentro de las dimensiones de Planck, la distinción real entre pasado y futuro se torna borrosa. Por analogía con el efecto túnel de los sistemas atómicos, que permite el paso de un electrón a través de una barrera de energía que no puede remontar, debemos esperar que en las distancias de Planck aparezcan procesos no permitidos en la teoría clásica de Einstein, incluyendo recorridos a velocidad superior a la de la luz. No sabemos cómo calcular las probabilidades de tales sucesos. En muchos casos, desconocemos hasta qué problemas hemos de resolver. No existen experimentos para guiarnos. Todavía es posible, pues, dar rienda suelta a nuestra imaginación.

Una de estas fantasías, persistente y repetidamente explicada en la bibliografía sobre gravedad cuántica, es la noción de topología fluctuante. El



8. VACIO CUANTICO, según lo imaginara John Archibald Wheeler, en 1957; presenta un aspecto caótico cada vez mayor a medida que nos fijamos en regiones más pequeñas del espacio. A la escala del núcleo atómico (*arriba*) el espacio parece totalmente uniforme. Ofrece cierta rugosidad en la geometría (*centro*) cuando consideramos regiones de unos 10^{-30} centímetros. A la escala de la longitud de Planck, unas mil veces menor (*abajo*), la curvatura y la topología del espacio sufren fluctuaciones violentas.



9. ¿CUANTAS DIMENSIONES TIENE EL ESPACIO? Es una de las preguntas que surgen ante la posibilidad de que el espacio-tiempo encierre una topología complicada. La superficie aquí presentada es bidimensional, pero sus conexiones topológicas le confieren una apariencia de objeto tridimensional. Podría suceder que el espacio de tres dimensiones que percibimos a escala macroscópica tenga, en realidad, menos dimensiones y sea topológicamente complicado. *(Compárese con la figura siguiente.)*

supuesto fundamental, introducido por Wheeler en 1957, es el siguiente: las fluctuaciones del vacío del campo gravitatorio, como las de todos los otros campos, aumentan su magnitud para longitudes de onda cortas. Si los resultados usuales de los campos débiles se extrapolan al dominio de Planck, las fluctuaciones de la curvatura resultan tan violentas que parecen capaces de producir agujeros en el espacio-tiempo y, por tanto, de cambiar su topología. Wheeler imagina el vacío en un estado de perpetuo desorden, con agusanamientos (y estructuras más complicadas) del tamaño de Planck, en continua formación y destrucción. El desorden sólo se "ve" a la escala de Planck. En una visión más tosca, el espacio-tiempo seguirá pareciendo uniforme.

Podemos señalar una objeción inmediata: todo cambio topológico va necesariamente acompañado por una singularidad en la estructura causal del espacio-tiempo; nos hallamos, pues, ante la misma dificultad que aparece en la visión de Hawking de la desintegración de un agujero negro. Supongamos, empero, que el punto de vista de Wheeler sea correcto. Una de las primeras cuestiones que debemos plantearnos será: ¿en cuánto contribuyen las fluctuaciones topológicas a la energía del vacío y cómo afectan a la resistencia (en el nivel tosco) del espacio-tiempo a curvarse? Hasta ahora, nadie ha dado con una respuesta convincente; y ello porque no se ha obtenido una descripción coherente del propio proceso de transición topológica.

Para calibrar uno sólo de los obstáculos que aparecen al intentar construir esta descripción, consideremos el proceso detallado en la figura 7. El dibujo presenta dos aspectos del mismo suceso: un agusanamiento acaba de aplastarse dejando como residuos dos pseudópodos en un espacio simplemente conexo. En una representación el espacio se dibuja curvado, en la otra plano.

Ahora consideremos el proceso inverso: la formación de un agusanamiento. Si existe una probabilidad finita para que un agusanamiento desaparezca al aplastarse, hay también una probabilidad finita de que aparezca otro. Y ya tenemos una nueva dificultad. Dando la vuelta al tiempo, la ilustración representa dos pseudópodos que han crecido espontáneamente en el vacío cuántico. En una de las ilustraciones la posibilidad de que los dos pseudópodos se junten para formar un agusanamiento parece razonable; en la otra no. Y, sin embargo, la situación física en los dos dibujos es la misma. La formación del agusanamiento parece razonable, en un caso, porque los pseudópodos parecen cercanos. Pero la "cercanía" no constituye ninguna propiedad intrínseca de la distribución espacial, como muestra el otro dibujo. La noción de "cercanía" exige la existencia de un espacio de mayor dimensión donde esté inmerso el espacio-tiempo. Además, este espacio de mayor dimensión debe poseer unas propiedades físicas tales que los pseudópodos se comuniquen que están cerca uno de otro. Pero entonces el espacio-tiempo ya no es el universo. El universo es algo más. Si continuamos fieles al punto de vista de que las propiedades del universo son intrínsecas y no proceden de algo exterior, parece fuera de nuestro alcance una descripción coherente de las transiciones topológicas.

Otra dificultad que presentan las fluctuaciones topológicas: cambian la dimensión macroscópica del espacio. Si los agusanamientos pueden formarse espontáneamente, también ellos pueden desarrollar otros agusanamientos, y así indefinidamente. El espacio puede alcanzar una estructura tal que, aunque tridimensional en la escala de Planck, tenga, aparentemente, cuatro dimensiones a escalas superiores. Un ejemplo corriente de este proceso es la formación de espuma que, constituida totalmente por superficies de dimensión dos, presenta estructura tridimensional [*véase la figura 9*].

Debido a dificultades de este tipo, algunos físicos piensan que la descripción usual del espacio-tiempo como un continuo uniforme es falsa en el nivel de Planck, y debe reemplazarse por otra. Nunca ha estado muy claro en qué consiste esta otra descripción. Teniendo presente el éxito de la descripción continua en un rango de escalas de longi-

tud que abarca una diferencia de órdenes igual a diez elevado a cuarenta (sesenta, si suponemos que el posible fallo tiene lugar justo en el nivel de Planck), parecería lógico suponer que la descripción continua es válida en todos los órdenes y que las transiciones topológicas no existen, simplemente.

Aun cuando la topología del espacio no cambie, no tiene por qué ser sencilla, ni siquiera en el nivel microscópico. Cabe imaginar que el espacio posea estructura de espuma, formada desde el principio; en cuyo caso, su dimensión aparente superaría la real. Su dimensión aparente puede ser también menor que la real.

Esta última posibilidad fue propuesta en una teoría desarrollada por Theodor Kaluza, en 1921, y Oskar Klein, en 1926. En la teoría de Kaluza-Klein, el espacio consta de cuatro dimensiones y, de cinco, el espacio-tiempo. La razón de que el espacio *parezca* tridimensional reside en que una de sus dimensiones es cilíndrica, tal como en el universo antes descrito, pero con una importante diferencia: el perímetro del universo en la dirección cilíndrica no mide miles de millones de años-luz, sino contadas unidades de Planck (tal vez de 10 a 100). De aquí que, cuando un observador intenta penetrar en la cuarta dimensión de espacio, se encuentra de vuelta casi instantáneamente al punto de salida. De hecho, no tiene sentido hablar de ese intento, porque los mismos átomos de que está compuesto el observador son muchísimo mayores que la circunferencia del cilindro. La cuarta dimensión, como tal, es sencillamente inobservable.

Sin embargo, puede ponerse de manifiesto de otra forma: ¡como luz! Kaluza y Klein demostraron que, si tratamos matemáticamente su teoría pentadimensional del espacio-tiempo del mismo modo que se trata la teoría cuatridimensional del espacio-tiempo en el modelo de Einstein, su teoría es equivalente a la combinación de la teoría del electromagnetismo de Maxwell con la teoría de la gravitación de Einstein. Las componentes del campo electromagnético están implícitas en la ecuación que satisface la curvatura del espacio-tiempo. Kaluza y Klein encontraron, pues, la primera teoría acertada del campo unificado, teoría que da una explicación geométrica de la radiación electromagnética.

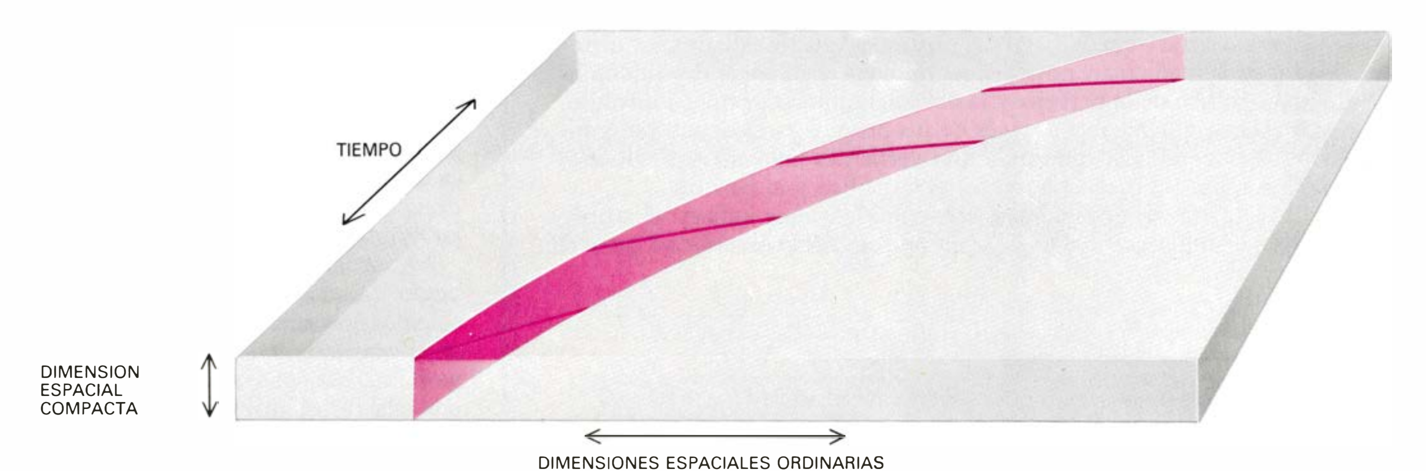
En cierto sentido, la teoría de Kaluza-Klein tuvo demasiado éxito. Aunque unió las teorías de Maxwell y de Einstein, no predijo nada nuevo y, por tanto, no pudo compararse con otras teorías. Ello se debió a que Kaluza y Klein pusieron, en la nueva dimensión, ligaduras al modo en que puede curvarse el espacio-tiempo. Si no fuera por estas ligaduras la teoría hubiese predicho nuevos efectos, pero esos efectos no parecían corresponder a la realidad. De aquí que, durante muchos años, se considerara su teoría como una bella curiosidad y se guardara en el armario.

La teoría de Kaluza-Klein fue liberada de su encierro en la década de los sesenta, cuando se puso de manifiesto que las nuevas teorías de aforo, a las que se prestaba cada vez más atención, podrían ser reformuladas como teorías del tipo Kaluza-Klein, en las que el espacio poseía varias dimensiones suplementarias a nivel microscópico, y no sólo una. Se empezó a pensar que toda la física podría explicarse en términos geométricos. Fue entonces cuando se planteó qué sucedería si se quitaban las ligaduras sobre las dimensiones compactas.

Una de las consecuencias es la predicción de fluctuaciones de la curvatura en las nuevas dimensiones; estas fluctuaciones se presentan como partículas con masa. Si la circunferencia de las nuevas dimensiones es de diez unidades de Planck, la masa de las partículas asociadas sería, aproximadamente, una décima parte de la masa de Planck o sea un microgramo. Como se precisa una enorme cantidad de energía para crear tales partículas, no se producirán casi nunca. En la práctica habrá, pues, pocas diferencias tanto si las ligaduras se imponen como si no. Quedan problemas; el principal es que la enorme curvatura de las nuevas dimensiones provoca una densidad muy alta de energía en el vacío clásico. Grandes energías para el vacío quedan eliminadas por el experimento.

Los modelos de Kaluza-Klein no han merecido todavía la debida atención; se desconoce qué papel puedan desempeñar en la física. Sin embargo, en los últimos dos o tres años han vuelto a ser analizados, esta vez relacionándolos con la notable generalización de la teoría de Einstein conocida como supergravedad, introducida, en 1976, por Daniel Z. Freedman, Peter van Nieuwenhuizen y Sergio Ferrara y (en una versión mejorada) por Stanley Deser y Bruno Zumino.

Uno de los inconvenientes de los modelos usuales de Kaluza-Klein es que predicen sólo la existencia de partículas con momento angular de espín igual a 0, $\hbar$ y $2\hbar$; estas partículas o no tienen



10. NUEVAS DIMENSIONES ESPACIALES, además de las tres ya conocidas, podrían existir si tienen forma "compacta". Una cuarta dimensión espacial podría estar enrollada como un cilindro de una circunferencia de unos 10^{-32} centímetros. En el dibujo, se ha "desenrollado" la dimensión compacta hipotética para que constituya el eje vertical de un diagrama del espacio-tiempo. La trayectoria de la partícula muestra así una componente cíclica; cada vez que alcanza la máxima magnitud de la dimensión compacta, se encuentra de nuevo en el punto de donde ha salido. La trayectoria que observamos es una proyección de la verdadera trayectoria sobre el espacio-tiempo macroscópico de cuatro dimensiones. Si la trayectoria es una geodésica, parecería el movimiento de una partícula cargada eléctricamente dentro de un campo eléctrico. Una teoría parecida fue propuesta en los años veinte por T. Kaluza y O. Klein, que demostraron cómo explicar juntos gravedad y electromagnetismo. Asistimos hoy a un renovado interés por estas teorías.

masa o la tienen muy grande. No admiten, pues, partículas de la materia ordinaria que, en su gran mayoría, poseen momento angular de espín igual a $1/2\,\hbar$. Sucede que, si reemplazamos la teoría de Einstein por la de la supergravedad y tratamos al espacio-tiempo como en la teoría de Kaluza-Klein, se produce una unión efectiva de todas las clases de espín.

En el "super"-modelo de Kaluza-Klein, más aceptado actualmente, se añaden siete dimensiones suplementarias al espacio-tiempo. Estas dimensiones tienen la topología de una esfera en siete dimensiones, espacio que, por sí mismo, posee algunas propiedades fascinantes. La teoría resultante, de extraordinaria riqueza y complejidad, incluye grandes supermultipletes de partículas. Las masas de las partículas son todavía o cero o enormemente grandes, pero es posible que, "rompiendo" la simetría de la esfera de siete dimensiones, se obtengan masas más realistas para algunas de las partículas. Persiste también la gran energía del vacío clásico, pero podría cancelarse con una energía negativa del vacío cuántico. No se sabe todavía si tendrán éxito estos esfuerzos por modificar la teoría. Queda, sin duda, mucho trabajo por delante antes de determinar con exactitud qué es lo que, en realidad, establece la teoría.

Si Einstein pudiera volver en espíritu a presenciar el curso que ha seguido su teoría, quedaría impresionado, agradablemente además. Le gustaría que los físicos al fin, tras años de duda, hayan aceptado su punto de vista de que las teorías que son matemáticamente elegantes deben ser estudiadas, aun cuando no parezcan adaptarse de modo inmediato a la realidad. Quedaría también satisfecho de que los físicos osen esperar la posibilidad de alcanzar una teoría de campos unificada. Estaría particularmente contento de que parezca empezar a ser verdad su viejo sueño de que toda la física pueda explicarse en términos geométricos.

Por encima de todo, quedaría asombrado. Asombrado de que la teoría cuántica persistiera entera e inamovible en el centro de todo, enriqueciendo la teoría de campos y siendo enriquecida por ella. Einstein nunca creyó que la teoría cuántica expresara una verdad definitiva. Nunca estuvo de acuerdo con el indeterminismo que de ella se deduce, y pensó que algún día sería sustituida por una teoría de campos no lineal. Ha ocurrido exactamente lo contrario. La teoría cuántica ha invadido la teoría de Einstein y la ha transformado.

Alternancia de presas en un ecosistema simple

En la isla de Terranova, pobre en especies de mamíferos, los linces hacían presa en la liebre americana hasta que la población de lepóridos se vino abajo. Cambiaron entonces a los terneros de caribú, mientras se recuperaba aquélla

Arthur T. Bergerud

La cadena ecológica a lo largo de la que se transfieren materia y energía desde los organismos fotosintéticos simples hasta los mamíferos carnívoros, pasando por los herbívoros intermedios, es muy compleja en la mayor parte del mundo. En cualquier zona restringida abundan las especies de presas y depredadores. Cada depredador se nutre de varios tipos de presa y las presas sufren el acoso de más de un depredador. Pero hay también regiones donde las condiciones geográficas han propiciado la formación de un sistema ecológico bastante simple que comprende sólo unas pocas especies. Tales ecosistemas constituyen un verdadero laboratorio para el estudio de los principios que rigen el tamaño de las poblaciones animales. La gran isla de Terranova, donde sólo existen 14 mamíferos indígenas, alberga uno de tales ecosistemas.

El conjunto de animales de Terranova, además de exiguo, está desequilibrado: nueve de las 14 especies son carnívoros. En una región con tantos carnívoros y tan pocos herbívoros de los que aquéllos puedan alimentarse, cualquier cambio de especie presa repercute de una manera importante; falta aquí la amortiguación de los efectos que se produciría en un ecosistema más complejo. Uno de esos casos espectaculares se observó en Terranova a principios de siglo, cuando el caribú y la liebre ártica, dos de las principales presas de la isla, empezaron a disminuir en número. Una investigación prolongada llevó a la conclusión de que, en ambos casos, la disminución se debía a un rápido aumento del número de linces.

Durante muchos siglos el lince escaseó en Terranova. Sin embargo, a raíz de la introducción en la isla de la liebre americana, o de pies blancos (*Lepus americanus*), en la década de 1860, la población de linces se expandió rápidamente, pues la aprovechó como fuente de alimento. En el primer tercio de nuestro siglo, las liebres americanas habían alcanzado su techo ecológico. Luego, la población de liebres americanas se "derrumbó", es decir, disminuyó rápidamente en número, dejando al lince sin su principal alimento. El lince, un depredador ingenioso, empezó entonces a hacer presa en los terneros de caribú y en la otra especie de liebre, la liebre ártica. Cuando la liebre americana se recuperó, el lince volvió a hacer presa en ella.

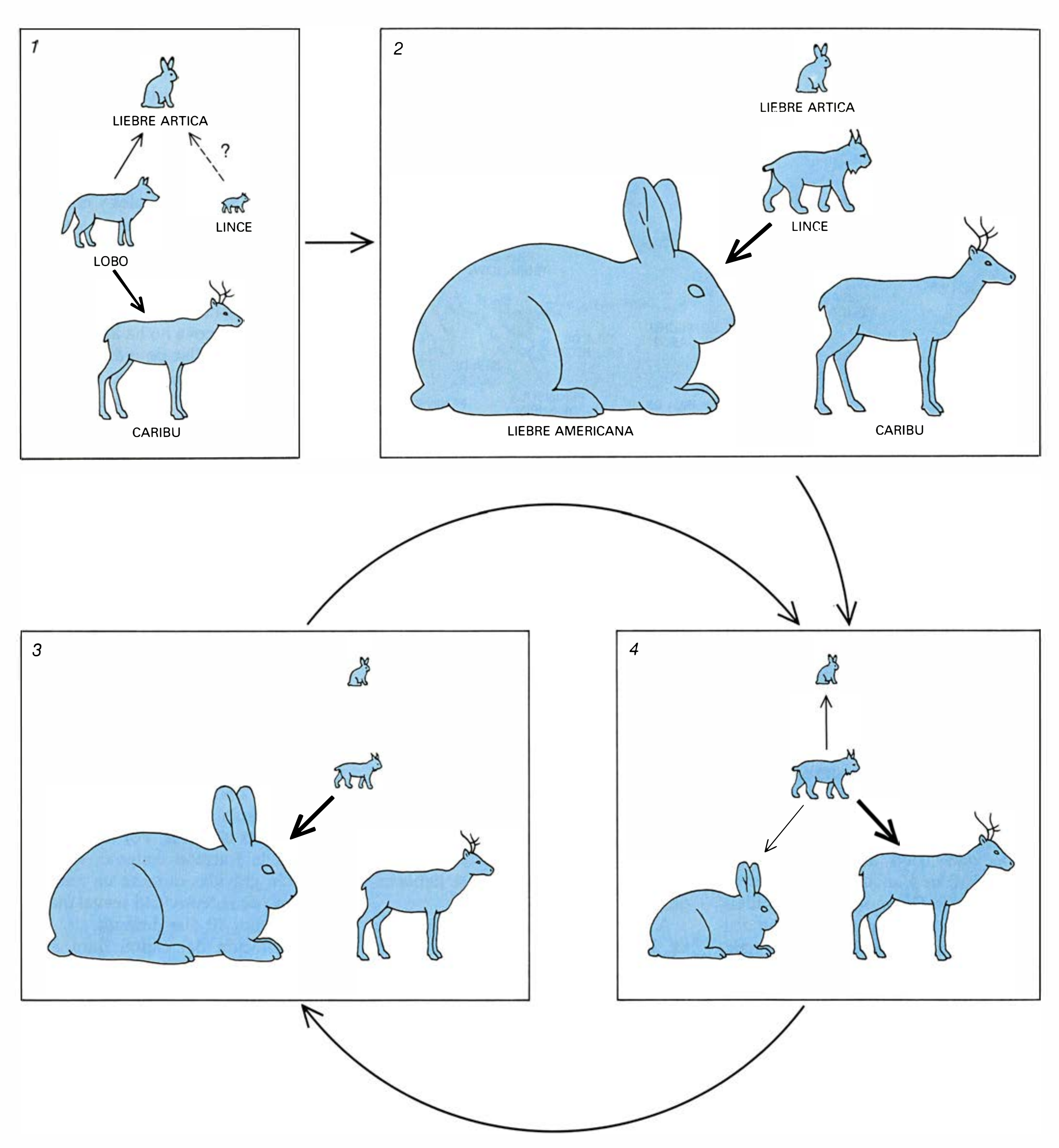
Desde entonces, el ciclo se ha repetido muchas veces. Así, en los últimos años, el lince, la liebre americana y el caribú constituyen los vértices de un triángulo en el que las poblaciones de las dos especies presa oscilan en ciclos alternos de aumento y disminución. Los linces limitan también el número de liebres árticas, que antaño abundaron en Terranova y ahora escasean. La introducción de una nueva especie en el ecosistema simple de la isla ha supuesto una larga y enmarañada serie de consecuencias para el equilibrio del mismo. El desenredo de la maraña nos ilustrará, a buen seguro, la dinámica de las poblaciones animales en las redes ecológicas de urdimbre más fina.

El sistema ecológico de Terranova difiere notablemente del que nos ofrecen las zonas vecinas del Canadá continental. Varios factores han contribuido a este aislamiento. Existen pruebas de que cuando la placa de hielo de Wisconsin avanzó, hace unos 18.000 años, los hielos no cubrieron toda Terranova. Parte de la isla debió quedar libre y sirvió de refugio, que permitió la supervivencia de los mamíferos locales. En consecuencia, algunas de las especies nativas se conservaron y evolucionaron independientemente de los mamíferos del continente. De los 14 mamíferos indígenas de Terranova, nueve se distinguen suficientemente de los tipos de tierra firme para que se les clasifique como subespecies, lo que es prueba de miles de años de aislamiento genético.

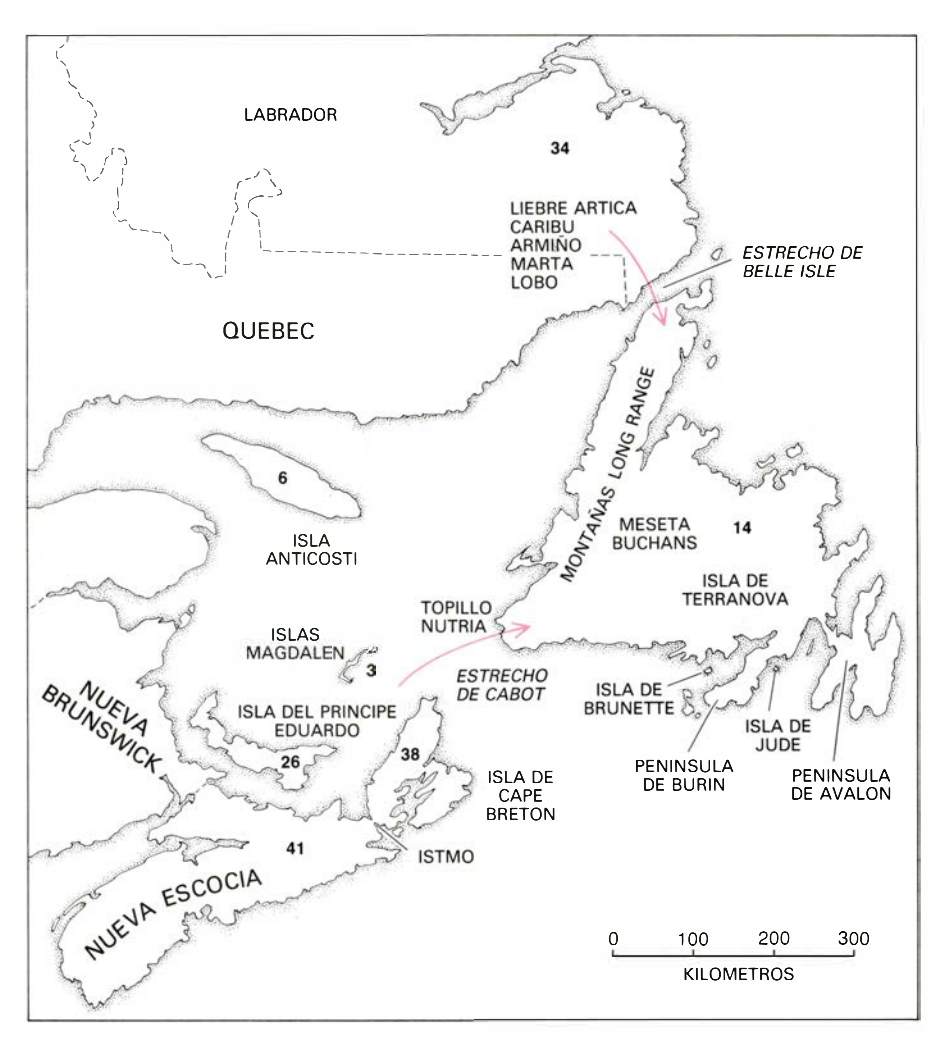
Algunas especies, en particular los depredadores, quizá llegaran a la isla sobre témpanos o puentes de hielo en el período postglacial. Sin embargo, el mar ha supuesto siempre una barrera al intercambio biológico entre la isla y las zonas costeras del Canadá continental. En Labrador, que dista sólo 17,6 kilómetros de Terranova a la altura del estrecho de Belle Isle, hay 34 especies de mamíferos. En la isla de Cape Breton, que se encuentra a 112 kilómetros al sudoeste de Terranova, y está conectada al continente, hay 38 especies.

El mar que rodea Terranova no ha actuado sólo de barrera, reduciendo el número total de especies, sino también de filtro, seleccionando determinados tipos de mamíferos y evitando que otros alcanzaran la isla. Los depredadores aventurados tenían más probabilidades que sus presas de colonizar la isla. Originariamente había siete cuadrúpedos carnívoros: lobo (*Canis lupus*), zorro (*Vulpes vulpes*), lince (*Lynx lynx*), nutria del Canadá (*Lutra canadensis*), oso negro (*Ursus americanus*), armiño (*Mustela erminea*) y marta (*Martes americana*). Los cinco herbívoros de los que estos depredadores se alimentaban antes de la llegada de la liebre americana eran el castor americano (*Castor canadensis*), la ratilla campestre americana (*Microtus pennsylvanicus*), el caribú (*Rangifer tarandus*, llamado reno en Europa), la liebre ártica (*Lepus arcticus*) y la rata almizclada (*Ondatra zibethica*). Hay también dos especies de murciélagos.

En este conjunto, la habitual pirámide de especies se halla invertida. Como todos los demás sistemas de transferencia de energía, los biológicos tienen una eficacia limitada; la masa total de depredadores es mucho menor que la de presas. Por la misma razón, el número de especies de depredadores es, por lo general, inferior al número de especies de presas. Las faunas invertidas simples, como la de Terranova, son frágiles, en parte porque constituyen una excepción a las reglas que normalmente rigen las cadenas alimentarias. Mi propia experiencia en la isla me confirma cuán vulnerable puede ser un sistema como éste. Desde 1956 hasta 1967 trabajé para la División de la Naturaleza de Terranova. Se me había contratado para descubrir la causa de un acusado descenso en el número de caribúes. Lo



1. CICLO DE ALTERNANCIA DE PRESAS en Terranova, ilustrado de forma esquemática. Durante milenios, y hasta mediados del siglo XIX, el principal depredador del caribú fue el lobo. Había pocos linces nativos (*1*). En 1864 se soltaron liebres americanas en la isla para alimento de pescadores. Se multiplicaron rápidamente, alcanzando un máximo hacia 1900 (*2*). Los linces, que comían liebres americanas, se multiplicaron también con rapidez. El lobo se extinguió hacia 1911. Cuatro años más tarde, la población de liebre americana se desmoronó y los linces se aprovecharon entonces de los terneros de caribú y de las liebres árticas, reduciendo considerablemente el número de ambas especies presa (*3*). Mientras tanto, las liebres americanas habían comenzado un ciclo intrínseco de 10 años de crecimiento y disminución de su población. Cuando la población de liebres alcanzó su máximo, los linces la explotaron de nuevo, permitiendo rehacerse a los caribúes (*4*). Durante décadas el lince ha alternado entre el caribú y la liebre americana en ciclos de una década. Las liebres árticas son hoy escasas; viven principalmente confinadas en regiones montañosas y su población fluctúa menos que las demás.



2. FAUNA ORIGINAL de la isla de Terranova, que incluía únicamente 14 especies de mamíferos. Las regiones vecinas del Canadá continental tenían muchas especies más. Cinco de Terranova pudieron llegar desde Labrador sobre témpanos o puentes de hielo; dos quizá procedan de las Provincias Marítimas. El mar ha significado una barrera para las especies y un filtro para los tipos de animales que alcanzaron la isla; a consecuencia de lo cual, la pirámide de la cadena alimentaria de las especies de Terranova aparece invertida: nueve de las 14 especies son carnívoras y sólo cinco son herbívoras. Tal tipo de fauna es frágil.

que encontré afectaba asimismo a otras especies.

En 1900, cuando el interior de Terranova se hallaba deshabitado, salvo por la pequeña tribu de indios micmac, abundaban las manadas de caribúes. A. A. Radclyffe Dugmore y J. G. Millais, dos de los primeros naturalistas en ocuparse del tema, estimaron que, a finales del siglo XIX, el número de caribúes de la isla rondaba las 150.000 o 200.000 cabezas. Quizá sus cálculos sólo fueran conjeturas teóricas. Ni recontaron ni muestrearon de forma sistemática las manadas. Sobre la base de entrevistas a cazadores más viejos, registros de caribúes cobrados y extensión del terreno a su disposición, he estimado que la población de caribúes en Terranova, en 1900, era de 40.000, lo que sigue siendo un número considerable de animales para una isla con una superficie de sólo 100.000 kilómetros cuadrados.

Veinticinco años después, el caribú se había extinguido casi por completo. Hacia 1925 Dugmore calculaba que sólo quedaban 200 cabezas. Tamaña disminución tuvo lugar a pesar de que el lobo, durante milenios fue el principal depredador del caribú, se había extinguido hacia 1911.

A finales del siglo XIX y principios del XX, la población nativa de liebre ártica redujo también sus proporciones primitivas, aunque ni el número original de liebres árticas ni su tasa de disminución están tan bien documentados como los del caribú. Sólo puede conjeturarse cuántas vivían en Terranova antes de que la isla se poblara intensamente, pero la mayoría de observadores se hallan de acuerdo en que el territorio que antaño ocupaban las liebres árticas se extendía desde la península Northern hasta la meseta Buchans y las montañas Long Range meridionales. Desde ahí, el área de distribución se extendía hacia el este, a lo largo de la costa sur hasta las penínsulas de Burin y Avalon. Este territorio incluye buena parte de la superficie de la isla.

El síntoma más claro de una reducción importante de las liebres árticas es la disminución de su área de distribución desde 1900. Outram Bang, un naturalista que describió siete de las 14 especies nativas de Terranova, anotaba en 1913: "Hoy, la liebre ártica es muy escasa en Terranova, isla a la que se circunscribe; vive en las cimas de las altas montañas". Hacia la década de 1950 las liebres árticas se hallaban confinadas en las tierras altas de las montañas Long Range septentrionales y meridionales y en la meseta Buchans, y se extendían a lo largo de la costa sur hasta la bahía d'Espoir. Incluso en esa zona escaseaba su número. En las últimas décadas, la población total de liebres árticas de Terranova no ha llegado probablemente al millar de individuos.

Cuando arribé a la isla se apreciaba mejor la disminución del caribú que la de las liebres árticas. La investigación sobre la disminución del caribú se concentró inicialmente en los factores que influyen en la tasa de mortalidad de sus terneros. En muchas especies de mamíferos, las clases de edad jóvenes constituyen el eslabón más débil en la supervivencia de la población. De hecho, la reducción en los rebaños de caribúes parecía deberse, sobre todo, a un aumento de la tasa de mortalidad entre los terneros.

Los machos adultos de la manada de caribú pasan buena parte del año forrajeando, separados de las hembras. A mediados de octubre, se reúnen durante el período de apareamiento, que comprende unos 10 días. En ese intervalo quedarán fecundadas el 80 por ciento de las hembras. Por lo que concierne a la fracción restante, algunas quedarán grávidas durante un período más corto de receptividad sexual que se produce unos 10 días después.

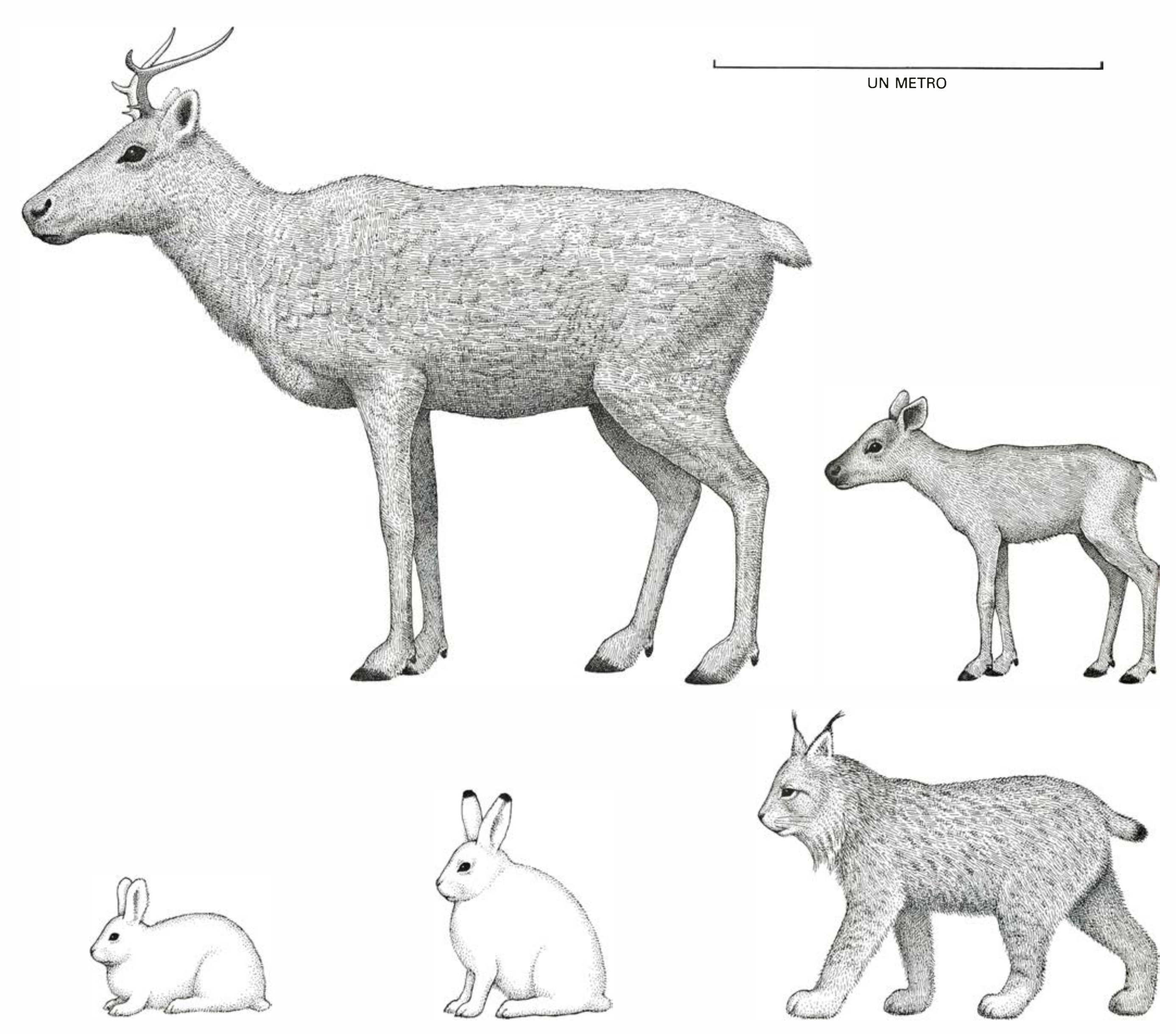
La gestación del caribú dura unos 229 días. Dada la sincronización del apareamiento, casi todos los terneros nacen en un período de 14 días, que empieza alrededor del 24 de mayo. Las hembras paren en grandes grupos en lugares específicos; suelen retornar cada año a la misma paridera. La mayoría de parideras de Terranova son zonas abiertas y pantanosas rodeadas de bosque achaparrado. Tras la época de nacimientos, algunas manadas se dispersan en los bosques para librarse de los insectos voladores.

Cada primavera, desde 1957 hasta 1964, los biólogos encontraban muchos

terneros muertos en las parideras. Era fácil localizar los restos porque la madre permanecía varios días cerca de su cría muerta. Muchos de los terneros muertos presentaban abscesos en el cuello, debidos a la infección de una bacteria patógena común: *Pasteurella multocida* tipo mucoide A. La causa inmediata de la muerte era una septicemia, o envenenamiento de la sangre, debida a la infección.

Puesto que las hembras de caribú paren en un ciclo anual bien sincronizado y la fracción de hembras que alumbran es relativamente constante de un año a otro, no es difícil calcular el número de nuevos miembros que se añadirían a un rebaño de caribúes en ausencia de influencias perturbadoras. Sobre la base de estos cálculos parecía que la tasa de mortalidad en los terneros era muy alta y que las crías halladas muertas representaban sólo una minoría de las que habían caído. De 1958 a 1967, un 27 por ciento de los terneros, por término medio, desapareció durante las dos semanas siguientes a su nacimiento. Mediante el recuento de las hembras con y sin terneros se estimó que, hacia octubre, al 70 por ciento, aproximadamente, de las hembras que habían parido les faltaba su hijo. Sin embargo, pasado octubre la tasa de mortalidad se reducía drásticamente; la mayoría de terneros que conseguían llegar a octubre sobrevivían al invierno. Lo que limitaba el tamaño de las manadas de caribúes era la mortalidad en los primeros cuatro meses de vida.

Además de la desaparición de terneros y de la concentración de las muertes en primavera, se planteaban otras cuestiones. En primer lugar, la proporción entre terneros macho y terneros hembra era, al nacer, de 52 a 48. En otoño la proporción se había invertido: 62 hembras por cada 38 machos. En segundo lugar, la mortalidad total era mayor en las parideras situadas a poca altitud. Lo que constituía un dato inesperado, porque en estas zonas la influencia del frío provocado por el viento era mínima. En las zonas alpinas, con un clima mucho más riguroso, la tasa de mortalidad descendía. Más intrigante resultó que, al contar las ma-



3. CUATRO ESPECIES DEL CICLO DE ALTERNANCIA DE PRESAS de Terranova. El caribú (*Rangifer tarandus*) tiene un ciclo de nacimientos sincronizado. Los terneros nacen durante un período de dos semanas en mayo y junio en parideras establecidas desde hace mucho tiempo. En la década de 1950 y de 1960 se encontraron muertas muchas crías en las parideras, con abscesos en el cuello. Las principales defensas de la liebre ártica (*Lepus arcticus; abajo, centro*) y de la liebre americana o de pies blancos (*Lepus americanus; abajo, izquierda*) son la vigilancia y la rapidez. La liebre americana vive en el bosque boreal, que en invierno se cubre de nieve blanda. La liebre ártica habita la tundra barrida por el viento, que tiene nieve dura. Ambas liebres se ilustran en su coloración invernal. El lince (*Lynx lynx*) acecha sus presas desde la maleza, las sujeta por el cuello y las arrastra bajo la vegetación.

nadas, apareció una pauta cíclica de supervivencia. Los censos anuales mostraban que la fracción de terneros que sobrevivía aumentaba en 1958 y 1959, y luego se reducía de manera constante de 1960 a 1963.

Las piezas del rompecabezas comenzaron a encajar en 1964, cuando se descubrió la causa de los abscesos. En la primavera de ese año se encontró un ternero muerto con cuatro heridas incisas en el cuello, pero sin abscesos. Parecía probable que los otros terneros muertos presentaran también incisiones, enmascaradas por los abscesos.

La fauna limitada de Terranova simplificó bastante la búsqueda de la causa de las perforaciones. Entre los depredadores locales solamente los caninos del lince tenían el tamaño y la disposición adecuados para producir las cuatro heridas uniformemente espaciadas. Cuando nos hicimos con un cráneo de lince para la verificación, se comprobó que los caninos del carnívoro se correspondían exactamente con las perforaciones.

Se cultivó saliva de la boca de un lince y se aisló de ella *Pasteurella multocida*. Encerramos entonces tres terneros de caribú en un redil con un lince, que mordió a los tres en el cuello. Separamos el lince de sus presas para que no las matara y sacamos a los terneros del corral. Pronto se desarrollaron abscesos con infecciones de *Pasteurella* y los terneros murieron, respectivamente, cuatro, cinco y quince días después de la mordedura.

Esta tarea detectivesca permitió concluir que la depredación por parte de los linces era lo que causaba el agotamiento de las manadas de caribúes. Los terneros muertos que se encontraron en el campo habían logrado escapar tras la agresión, pero acabaron muriendo de septicemia. Los biólogos exploraron las zonas forestadas en los bordes de las parideras y descubrieron restos de otras crías.

Depredadores felinos, los linces muerden a su presa por el cuello y la arrastran bajo el dosel de la vegetación baja. Aparentemente, el lince acecha bajo los matorrales en el lindero de la paridera, a la espera de que se acerque un ternero incauto. Sale entonces disparado, muerde a la cría en el cuello y lo arrastra hasta la maleza. Los terneros macho son más vulnerables al ataque que las hembras; los primeros, más curiosos, a menudo se apartan de la madre para explorar el lindero del bosque. Las hembras tienden a permanecer cerca de la madre. La hembra adulta de caribú en Terranova mide alrededor de 1,2 metros a la cruz y pesa unos 100 kilogramos, tamaño más que suficiente para ahuyentar a un lince atacante, que pesa unos 10 kilogramos.

Para comprobar la hipótesis de que el lince era el responsable de la disminución de las manadas de caribúes se eliminaron los linces, mediante el uso de trampas, de la zona que rodeaba las dos parideras. Se apreció entonces un aumento significativo de la fracción de terneros que sobrevivían hasta unirse a la manada. También se introdujeron caribúes en varias pequeñas islas de la costa de Terranova donde no había linces. Para control del experimento se introdujeron caribúes en hábitats continentales donde se habían extinguido los caribúes locales, pero había linces. En las pequeñas islas el caribú aumentó a una tasa cercana a la máxima calculada para la especie. En los hábitats donde había linces la tasa de aumento fue la mitad de la máxima. En ese mismo período el tamaño de la manada de Terranova cambió muy poco. Se componía de 6100 animales en 1961 y de 6200 en 1966.

Uno de los lugares en los que el caribú había aumentado rápidamente desde su introducción era la isla de Jude. De la repoblación se encargó el Servicio de la Naturaleza de Terranova, bajo la dirección de Eugene Mercer. El propio Mercer participó activamente en los trabajos de Jude. Se fijaron radiotransmisores a dos linces y se les soltó para observar su efecto sobre el rebaño de caribúes en un ambiente antes libre de depredadores. En la primera temporada, el 65 por ciento de los terneros cayó presa de los linces, mientras que en el año anterior todos los terneros habían sobrevivido.

El estudio de la isla de Jude reveló también la razón por la que la madre permanece cerca del cuerpo de su ternero durante varios días. Los movimientos de los linces, seguidos por las transmisiones de radio, indicaron claramente que después de haber sido ahuyentado de la cercanía de su presa por la madre caribú el lince solía permanecer escondido en la maleza inmediata, a la espera de que aquélla se alejara. Cuando se retiraron de la isla los linces portadores de transmisores, la tasa de superviviencia de terneros recuperó inmediatamente su elevado nivel anterior.

Después del experimento de Jude, hasta los más escépticos quedaron convencidos de que el lince, que nunca se había tenido por depredador importante del caribú en el Canadá continental, constituía el principal factor limitante de la población de caribúes de Terranova. El lobo, antaño la mayor amenaza a la población de caribúes, había desaparecido, y el lince había ocupado su lugar en el ecosistema simple de la isla.



4. LA CAUSA DE LOS ABSCESOS en los terneros de caribú muertos se descubrió en 1964. Se encontró el cuerpo de una cría con cuatro heridas punzantes en el cuello, pero sin abscesos. Los cuatro grandes caninos de un cráneo de lince ajustaban claramente en las heridas. Al cultivar saliva de un lince se aislaron las bacterias patógenas que se habían recuperado antes en los abscesos. Se probaba así que la depredación por los linces era responsable de la fuerte mortalidad observada entre los terneros de caribú.

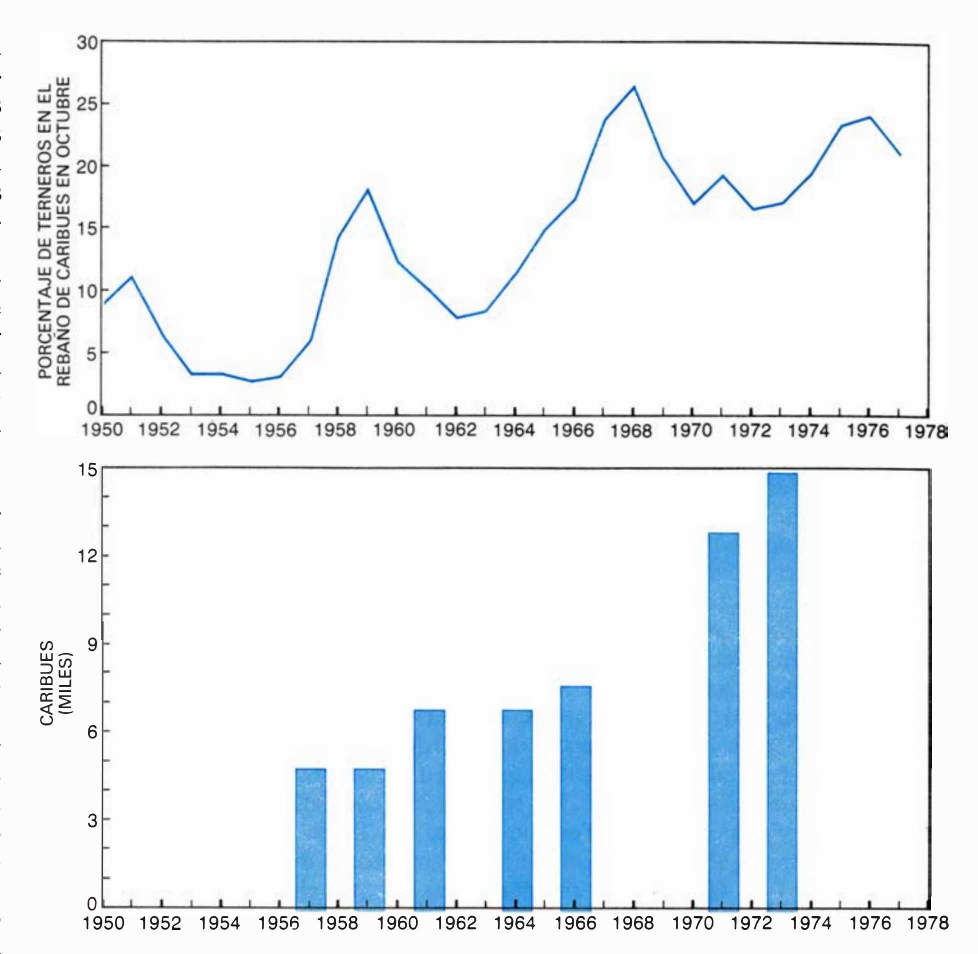
Si el lince podía regular la población de caribúes, parecía verosímil que regulara asimismo el número de liebres árticas. Las dos parideras de las tierras altas con mayor tasa de supervivencia de los terneros eran asimismo las zonas donde ocasionalmente se señalaba todavía la presencia de liebres árticas. Por el contrario, quizá las liebres americanas habían reducido la población de liebres nativas al competir con ellas por el alimento. Un cambio de la flora de la isla que eliminara los recursos alimentarios de la liebre indígena explicaría igualmente su escasez.

Para determinar si el aumento de liebres árticas encontraba un límite en la cantidad de alimento disponible, se liberaron dos machos y dos hembras en la isla de Brunette, situada unos 16 kilómetros al sur de Terranova, en la bahía de Fortune; no había allí mamíferos depredadores ni liebres americanas. Más aún, como buena parte de Terranova, la isla de Brunette era un hábitat adecuado para las liebres árticas. Gran parte de la isla está ocupada por zonas alpinas y subalpinas, y las cumbres de tundra de las colinas se tapizan de musgos y matorrales de ericáceas. Al cabo de seis temporadas reproductoras había 1000 liebres árticas en la isla. Al igual que el caribú, en ausencia de los linces las liebres se habían multiplicado a una tasa cercana a la máxima de la especie.

La suelta en Brunette probó que el alimento no suponía una barrera para el crecimiento de la población de liebre ártica. A continuación se utilizaron liebres árticas de Brunette para comprobar las otras hipótesis; se introdujeron en islas donde había liebres americanas pero no linces. Como control del experimento se soltaron liebres árticas en zonas de Terranova en las que había liebres americanas y linces.

En las islas externas, donde habitaban liebres americanas pero no linces, las liebres árticas sobrevivieron y su población aumentó, aunque no tan rápidamente como en la isla de Brunette. En los hábitats de Terranova, con linces, ninguna de las introducciones tuvo éxito. Así pues, la liebre indígena podía soportar la competencia de la liebre de pies blancos, pero no la depredación del lince, al menos en las tierras bajas de Terranova.

Los experimentos de campo confirmaron que el lince limitaba la abundancia del caribú y de la liebre nativa en Terranova. A partir de los experimentos de suelta y del estudio de los registros históricos sobre la fauna de la isla, empezaba a perfilarse un insólito sistema dinámico que comprendía el lince, el caribú y las dos especies de liebres. Durante miles de años, antes de que el hombre blanco poblara Terranova, las liebres árticas, los lobos y los caribúes coexistieron en los yermos y macizos de la isla azotados por el viento. En ese primer período el lince era poco común. Tanto, que los expertos no se ponen de acuerdo sobre si debe siquiera considerársele nativo de la isla.



5. TASA DE SUPERVIVENCIA de los terneros de caribú, que determinó el tamaño del principal rebaño de estos animales en las décadas de 1950 y 1960. La curva superior muestra qué fracción del rebaño principal correspondía a terneros en octubre. Recoge el número de terneros que sobrevivieron a la primavera y el verano. Los picos de supervivencia de terneros en 1951, 1958, 1968 y 1976 coincidieron con máximos de liebres americanas en la isla. Cuando la población de liebres caía, se reducía la supervivencia de los terneros. Los histogramas del cuadro inferior indican el tamaño total del rebaño en el interior de Terranova en los años para los que se dispone de datos. Cuando el ciclo de cambio de presas era más intenso, el tamaño de la manada se determinaba por la supervivencia de los terneros. Debido a la disminución de los linces y a la introducción de nuevas especies, la intensidad de los ciclos ha bajado.

Es indudable que las poblaciones de lobos, caribúes, linces y liebres árticas fluctuaban, pero existían mecanismos estabilizadores que evitaban que se produjera la extinción de ninguna de las especies. En 1864 se llevaron liebres americanas a Terranova para ayudar a la alimentación de pescadores aislados en la costa occidental. La introducción del nuevo tipo de liebre trastocó la delicada dinámica de poblaciones que se había establecido entre las especies indígenas.

Cuando un animal como la liebre americana llega a un ambiente nuevo, con frecuencia tiende a aumentar rápidamente en número. Pasadas unas pocas décadas, la población introducida alcanza su nivel máximo. Hacia 1890 toda la isla de Terranova experimentaba ese "máximo" en la población de liebres de pies blancos. La densidad de los animales era tan elevada que algunos habitantes podían atrapar hasta 1000 liebres en una sola temporada.

Tras el máximo inicial que sigue a la introducción de una especie suele darse una disminución de la población, a medida que los límites intrínsecos del nicho ecológico ocupado por la especie comienzan a ejercer sus efectos. Hacia 1915, la población de liebre americana comenzó a disminuir rápidamente.

En gran parte del Canadá el lince es el principal depredador de la liebre americana. Al tiempo que, hacia finales del siglo XIX, la población de liebres americanas de Terranova crecía rápida-

mente, la población de linces también aumentaba, aunque la curva del aumento de los linces aparecía retrasada con respecto a la de las liebres. Es esta una pauta típica en las zonas donde el lince y la liebre americana forman una pareja obligada depredador-presa. En tales zonas, las curvas que representan las dos poblaciones adoptan una forma sinusoidal regular. La población de liebres americanas crece hasta que se ve limitada por el alimento o por un mecanismo biológico intrínseco. Después, la población de liebres disminuye, pues su tasa de reproducción se reduce. Al decaer el número de liebres, baja también la tasa de reproducción de los linces y mengua la población de linces. En ausencia de perturbaciones externas, el ciclo completo dura de nueve a 11 años.

A principios de siglo el lince se había extendido por toda Terranova, procedente de la costa occidental. Se avistaron linces dentro de los límites urbanos de la ciudad de St. John's, la capital de la península de Avalon, que se extiende desde la costa sudoriental. En 1904 se propuso ante la Asamblea de Terranova un proyecto de ley para exterminar a los linces, lo que da idea de la rapidez con que estos animales habían aumentado. Sin embargo, la población de linces se regularía en último término por la incidencia de factores ecológicos y no por factores humanos.

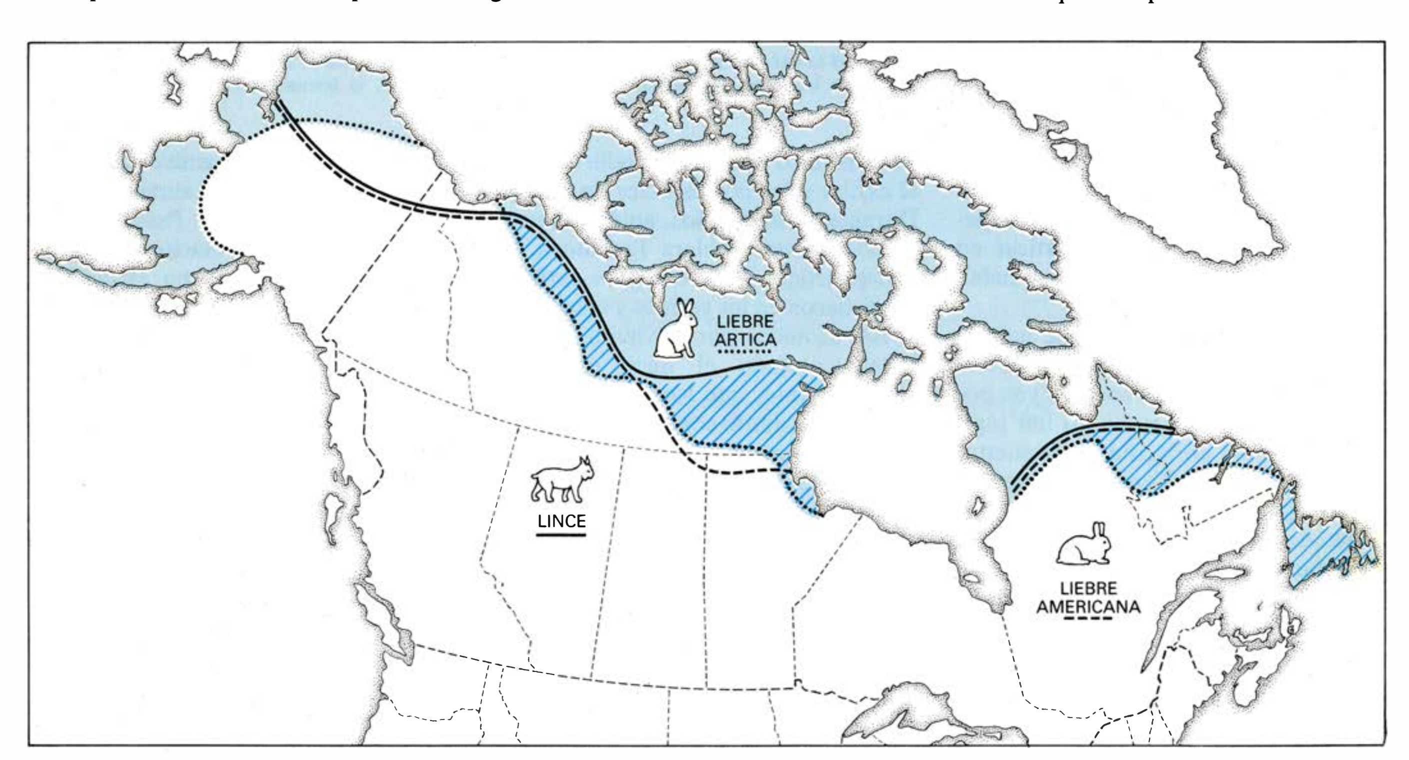
La población de liebres de pies blancos permaneció en un máximo desde 1896 hasta 1915. (No se habían producido aún las oscilaciones de las poblaciones de linces y liebres.) Cuando la población de liebres se desplomó, los abundantes linces se quedaron sin su principal alimento. Los linces son carroñeros ingeniosos y depredadores flexibles. Cuando el alimento es particularmente escaso, los adultos dejan de reproducirse, pero rara vez mueren de hambre. Utilizan especies presa salvajes alternativas, se dirigen a los animales domésticos, o bien a desperdicios oceánicos depositados en la costa.

En Terranova, las únicas presas salvajes alternativas eran ratones, liebres árticas y terneros de caribú. Presumiblemente los linces empezaron a cazar al acecho terneros de caribú y las pocas liebres árticas disponibles. No bastaban las nuevas presas para permitir que el número de linces creciera o se mantuviera estacionario, pero mitigaron la pérdida de la principal fuente de alimento.

Casi coincidiendo con la época en que los linces empezaron a capturar gran número de terneros de caribú, la tasa de mortalidad entre los caribúes adultos aumentó por una razón completamente distinta. El ferrocarril de Terranova, terminado hacia 1900, corta la ruta que seguía antaño la principal manada de caribúes de la isla en su migración otoñal hacia el sur. Desde 1911 hasta 1925, los cazadores, que utilizaban el ferrocarril para perseguir a los animales, mataron sin tiento. En 1911, 1912, 1914 y 1915, los cazadores cobraron anualmente más de 5000 caribúes.

Puesto que la gran reducción en los caribúes adultos se produjo cuando los linces hicieron presa en sus terneros, faltaron nuevas generaciones que sustituyeran las pérdidas causadas por los cazadores. En unos pocos años se exterminaron prácticamente los 40.000 caribúes que había en 1900. En las dos décadas que se precisaron para la reducción catastrófica de la manada de caribúes la población de liebres americanas comenzó a fluctuar con la pauta cíclica observada en otras partes del Canadá. Se observaron picos en el número de liebres americanas en 1920, 1931, 1940-43, 1951-52, 1959-60, 1969 y 1976.

En los picos del ciclo de la liebre de pies blancos los linces volvieron a su presa principal y las tasas de supervivencia de los terneros de caribú aumentaron. Los registros históricos muestran que los rebaños crecieron rápidamente hacia 1940, 1950 y 1960. Durante el máximo de liebres americanas de 1960, los principales rebaños de caribúes que se contaron aumentaron desde 4800 cabezas, en 1959, hasta 6100, en 1961. Así, la adaptación depredadora del lince significaba que las poblaciones de cari-



6. AREA DE DISTRIBUCION DE LA LIEBRE ARTICA en Canadá (*color*); podría estar limitada, en parte, por relaciones depredador-presa. El límite meridional del área de distribución de la liebre ártica coincide con el límite septentrional del área de distribución del lince: ambas áreas se solapan sólo ligeramente (*región rayada*). Sin embargo, las áreas de distribución del lince y de la liebre americana se superponen en gran medida. El lince y la liebre de pies blancos tienen una carga pedia baja: la presión que ejercen los pies sobre el suelo cuando el animal corre es pequeña. Por ello están bien equipados para la carrera sobre la blanda nieve del bosque boreal. La liebre ártica tiene una carga pedia superior, para correr sobre la nieve dura de la tundra. En el bosque boreal, la liebre ártica tropieza y es vulnerable al lince, sin cuya presencia podría extenderse más al sur. Las ilustraciones son de Tom Trentiss.

búes y de liebres árticas quedaban estrechamente acopladas a la población de liebres americanas.

En las últimas décadas los ciclos de supervivencia de los terneros de caribú se han hecho menos pronunciados; incluso en los puntos bajos del ciclo sobrevive una fracción mayor de terneros que cuando el ciclo era mucho más intenso. Esta amortiguación es consecuencia, en parte, de que la población de linces de las islas se ha reducido, y ello debido a otro fenómeno cíclico: la moda femenina.

Después de un período de transición, las pieles de animales de pelo largo han vuelto recientemente a ponerse de moda. Resulta de nuevo rentable la caza del lince mediante trampas; y así, ha bajado la densidad de la población de linces en Terranova. Además, se han importado a la isla nuevos tipos de presa: el grévol engolado (*Bonasa umbellus*) y el grévol del Canadá (*Canachites canadensis*), entre otros. La cadena alimentaria se ha complicado, lo que representa un efecto amortiguador adicional a las oscilaciones de las poblaciones de depredadores.

En gran parte, el importante efecto de los linces sobre los caribúes de Terranova se explica por la larga adaptación del caribú a la depredación por parte del lobo. Las estrategias que daban resultado contra el lobo se tornan desventajosas frente a un nuevo depredador con costumbres de caza distintas. Por ejemplo, las grandes agregaciones de caribúes que se forman después de la cría en primavera proporcionan a estos animales una ventaja sustancial al proteger a los terneros de los lobos, que suelen cazar de día y acechan a su presa en terreno abierto. Cuando el rebaño de caribúes avista un lobo, su principal estrategia defensiva es la huida a gran velocidad. El tamaño del rebaño en la época de los partos mejora la vigilancia: aumenta la probabilidad de que una hembra aviste desde lejos al lobo que viene. Además, el ojo del caribú está bien adaptado a detectar el movimiento durante el día, un valioso don frente a un depredador diurno que opera en campo abierto. Por la misma razón, el terreno despejado de la paridera facilita la detección de un lobo al acecho.

Sin embargo, esta combinación de características fisiológicas y de comportamiento tiene poco valor cuando el depredador es un lince. El lince caza principalmente de noche; tiende una emboscada a su presa, en vez de derribarla después de una persecución a toda velocidad. Los linces suelen agazaparse y permanecer inmóviles durante su acercamiento, lo que neutraliza las defensas del caribú, cuyos ojos están mejor adaptados a la detección de movimientos que a la detección de formas. Por lo general, un caribú no se alarma ante la presencia de un hombre que permanece quieto; probablemente tampoco reaccione frente a un lince agazapado.

Cuando un depredador acecha a un rebaño de herbívoros, con frecuencia selecciona un animal determinado al que atacar porque muestra algún rasgo que lo distingue de la manada. Los lobos eligen a caribúes jóvenes, enfermos, viejos o lisiados; casi cualquier tipo de diferencia aparente puede hacer que un caribú llame la atención del lobo. En cambio, en los linces este tipo de selección se centra exclusivamente en los jóvenes. Para un lince el blanco más atrayente es un ternero curioso y activo en el lindero de la paridera, cerca del borde de vegetación que sirve de cobijo de acecho al depredador.

Quizá la depredación por parte de los linces, en ausencia de lobos, señale una nueva dirección en la evolución del caribú de Terranova. La selección natural podría ahora favorecer la dispersión de los terneros en lugar de su concentración en la paridera abierta, que les convierte en fácil blanco. Podrían estrecharse, además, los vínculos entre la madre y el ternero. Un ternero que permanece cerca de una madre protectora tiene más probabilidades de sobrevivir a la depredación por parte de un lince que un ternero que corra deprisa, aunque éste sea más capaz de escapar de un lobo.

Si el caribú se adapta, su tasa de supervivencia podría aumentar rápidamente, porque la adaptación recíproca del lince debe operar entre límites estrechos. El lince no puede depender únicamente del caribú para su alimento; vive sobre todo de la liebre americana. Cualquier cambio en su comportamiento de caza no debe alterar la pauta que le convierte en eficaz cazador de liebres. Una nueva relación evolutiva entre el lince y el caribú favorecería, por tanto, a este último.

Al igual que la depredación del lince sobre el caribú, la relación entre el lince y la liebre ártica no cuenta con un período prolongado de estrecha coevolución. Las dos especies se distribuyen en áreas que en su mayor parte se hallan separadas. Sobre un mapa del Canadá, la frontera meridional del área de distribución de la liebre ártica coincide exactamente con la frontera septentrional del área de distribución del lince. En cambio, los territorios del lince y de la liebre americana se superponen en su mayor parte: el depredador se extiende hacia el norte casi tanto como la presa.

La mayoría de los ecólogos explicaría la distribución de las tres especies según el concepto de biomas, o zonas biológicas: la liebre ártica vive en el bioma de la tundra, mientras que el lince y la liebre de pies blancos ocupan el bioma del bosque boreal. Sin embargo, las relaciones depredador-presa también pueden influir sobre el área de distribución de la liebre ártica. Quizás, en ausencia de linces, la liebre ártica se distribuiría en el Canadá más al sur del área que ocupa hoy.

El éxito del lince en su caza al acecho de las dos especies de liebre depende, entre otras cosas, de la vegetación y de las condiciones de nieve. En hábitats forestales el lince acecha y embosca a su presa; la liebre americana se sienta inmóvil y alerta. El resultado de un encuentro entre ambos animales depende principalmente de la distancia que los separa cuando la presa avista al depredador y sale corriendo.

Mercer ha señalado que, cuando el lince y la liebre de pies blancos corren sobre la nieve suelta del bosque boreal, ambos animales se hallan equiparados en un factor decisivo: su carga pedia baja; es decir, la presión que ejercen los pies sobre el suelo cuando los animales corren es pequeña. Así, pueden correr veloces por el bosque sin hundirse en la nieve blanda.

En cambio, la liebre ártica está adaptada a la nieve dura de la tundra septentrional. Tiene una carga pedia que dobla de lejos la característica de la liebre americana. Sobre nieve blanda la liebre ártica tropieza, con lo que se reducen mucho sus posibilidades de escapar de un lince. Ello no supone que las condiciones de la nieve por sí solas determinen el área de distribución de la liebre ártica. Es la presencia de nieve blanda y de depredadores con una carga pedia baja lo que constituye la combinación fatal.

La alternancia de los depredadores de una especie presa a otra delimita hoy un área de investigación particularmente prometedora en ecología. Los cambios ayudan a explicar de qué modo puede mantenerse la estabilidad en ecosistemas pequeños y frágiles. Cuando la presa primaria disminuye, un depredador ingenioso puede dirigirse a una presa secundaria. Mediante este mecanismo, el depredador mantiene su lugar en una red alimentaria que de otro modo se desmoronaría. El paso de una presa a otra también elimina la

presión de depredación de una población de presas en declive y le permite sobrevivir.

Incluso así, la alternancia de presas puede conducir a que la población de presas en declive sobreviva únicamente a una densidad muy baja. Esto es lo que ocurre con la liebre ártica en Terranova. La gran densidad de los linces, que se apoya en la liebre de pies blancos, asegura que la densidad de liebre ártica permanezca baja.

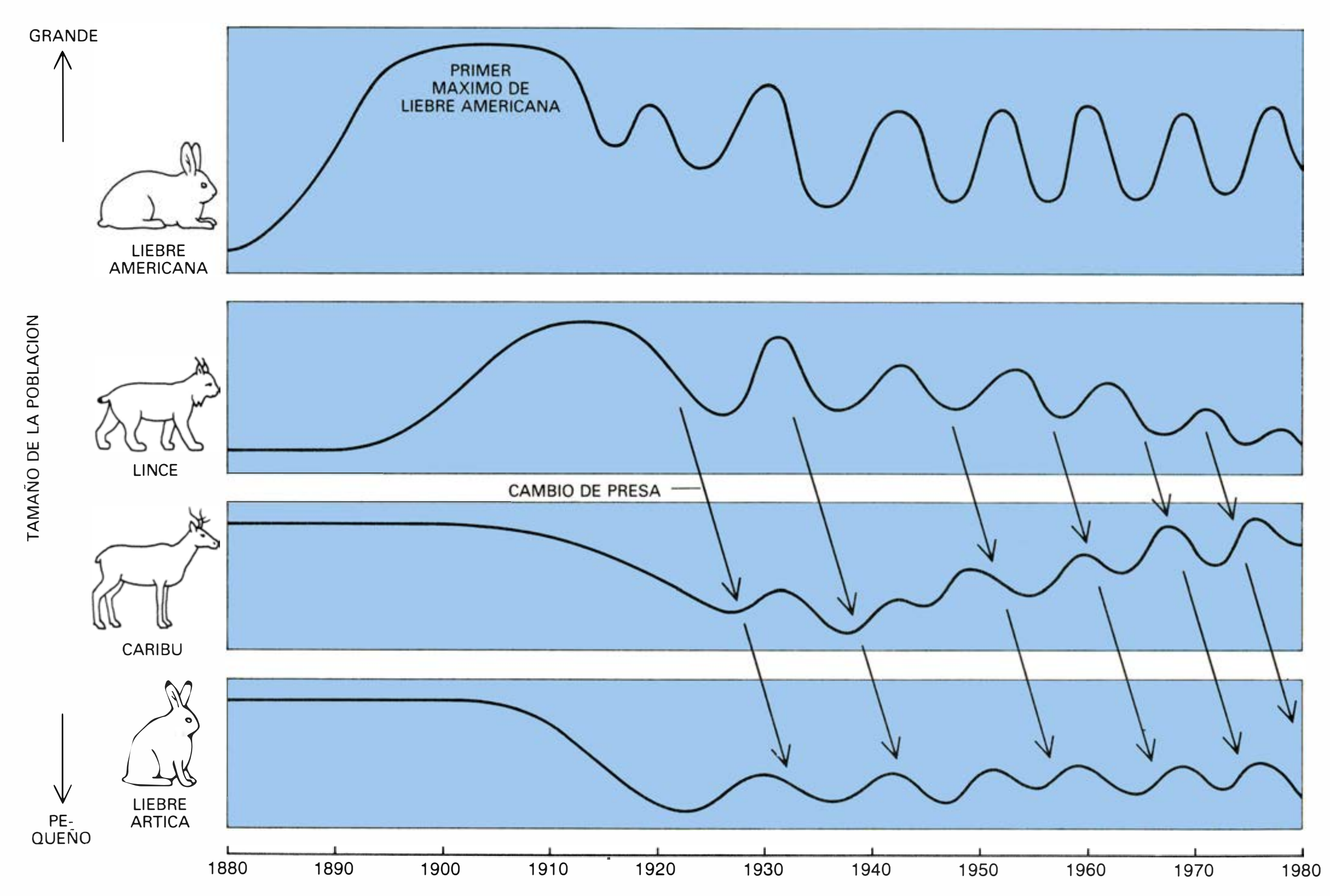
En un sistema de este tipo se dice que la liebre ártica se halla en el "pozo del depredador". Otro ejemplo de especie presa que se encuentra en esta situación es el caribú de Ontario y Columbia Británica. En estas provincias el caribú se asienta en un vértice de un sistema triangular cuyos otros ángulos ocupan el lobo y el alce. En ambas regiones, el caribú coevolucionó con el lobo en ausencia del alce. En el siglo XIX el alce (*Alces alces*) se extendió por estas provincias desde el sur y el este. La adición del alce a la base alimentaria hizo que la población de lobos prosperase. Sin embargo, a los lobos les es más fácil cazar caribúes que alces, y el lobo mata caribúes siempre que puede.

En ciertas partes de Ontario y de Columbia Británica los caribúes se hallan en el pozo del depredador. Persisten amparados en su escasez. Si aumentaran, los lobos caerían sobre ellos y bajaría su número. No pueden escapar del pozo mientras la población de alces continúe sosteniendo una elevada densidad de lobos.

En algunos casos, la especie presa de uno de esos sistemas sobrevive, además de porque su número de individuos es bajo, porque la población de dicha especie está tan esparcida que cazarla es poco eficaz en términos energéticos. Por ejemplo, la actual capacidad de aguante de Terranova para la liebre ártica viene definida por la superficie mínima por liebre que se precisa para hacer que los encuentros entre los linces y su presa sean muy infrecuentes.

De hecho, tales encuentros deben de ser lo suficientemente raros para que la supervivencia de las liebres árticas jóvenes compense la mortalidad entre los adultos. Si el número, y con ello la densidad de las liebres árticas, aumenta, se viola esta condición. La depredación por los linces restablecería rápidamente el equilibrio. Si ello es así, las relaciones espaciales resultarían una característica altamente significativa de la dinámica de los sistemas depredador-presa.

La conclusión más significativa que se extrae del estudio de la fauna de Terranova es, sin embargo, el efecto considerable e impredecible de introducir una nueva especie en un ecosistema simple. ¿Qué habitante de la isla habría imaginado en 1864 que la importación de unas pocas liebres americanas como fuente de alimento para unos pescadores hambrientos produciría, un siglo más tarde, ciclos en la supervivencia de los terneros de caribú y limitaría la población de liebres árticas? Cuando interfiere con los sistemas naturales insulares y frágiles, el hombre debe siempre proceder con cautela.



7. DINAMICA DE POBLACIONES de la liebre americana, el lince, el caribú y la liebre ártica; sugiere que la fuerza impulsora del sistema de alternancia de presas es el ciclo intrínseco de 10 años de la liebre americana. El primer máximo en la población de liebres americanas de Terranova fue el más alto; los picos subsiguientes bajaron algo. Antes de que la liebre americana llegara a la isla, los linces escaseaban tanto que se duda de que el animal sea en realidad una especie nativa. Cuando la población de liebres americanas creció, el número de linces aumentó rápidamente. Aunque se afirma que había unos 40.000 caribúes en 1900, a consecuencia de la depredación por linces y por cazadores humanos quedaban 1000 cabezas o menos hacia 1925. La población de la liebre ártica se halla en un "pozo del depredador": continúa existiendo amparada en su escasez. La especie no puede salir del pozo. Si el número de liebres árticas aumenta, los linces lo reducen rápidamente, pues las explotan al producirse una disminución cíclica de las liebres americanas.

Aerodinámica de los vehículos terrestres de propulsión humana

La resistencia al flujo del aire estorba a la bicicleta y al ciclista. Los diseños aerodinámicos han producido vehículos que van, sin ayuda, a más de 90 kilómetros por hora en llano

Albert C. Gross, Chester R. Kyle y Douglas J. Malewicki

Durante décadas se han aplicado, con gran éxito, los principios de la aerodinámica en la mejora de velocidad y eficiencia en aviación, automovilismo, motocicletas y hasta en esquí y patinaje de competición. Sin embargo, los vehículos impulsados por energía humana fueron prácticamente ignorados hasta hace poco, lo que resulta extraño si advertimos que la resistencia del aire es, con mucho, la mayor fuerza de frenado que les afecta. En una bicicleta, por ejemplo, a velocidades superiores a 30 kilómetros por hora, llega a más del 80 por ciento de la fuerza total de frenado. Vamos a explicar aquí este abandono y a presentar la atención que comienza a recibir la aerodinámica en el rendimiento de los vehículos terrestres de propulsión humana.

Fijémonos en la bicicleta. Su forma se ha mantenido casi igual durante cerca de un siglo. La Rover Safety Cycle, introducida en Inglaterra en 1884, podría pasar fácilmente por una bicicleta moderna; carece únicamente del tirante del asiento, que daría lugar al actual cuadro, y otros accesorios, como los frenos y los desarrollos múltiples. Los diseñadores y usuarios de bicicletas reconocieron, casi desde sus comienzos, la importancia de la aerodinámica, pero restricciones de diseño artificiales impidieron durante mucho tiempo la aplicación de la tecnología necesaria. Era tan obvio entonces como ahora que a velocidades de competición, entre 32 y 48 kilómetros por hora, las fuerzas del viento eran grandes.

Antes de 1900 se había hecho habitual la posición en cuclillas del corredor para reducir así la resistencia del aire. Otra práctica adoptada antes de 1900 fue la de poner un tándem múltiple delante del ciclista que lo protegiera del viento. En 1895, el rodador galés Jimmy Michael recorrió 46 kilómetros en una hora detrás de un tándem de cuatro. En 1899, Charles Murphy (alias "Milla por Minuto"), de Estados Unidos, adquirió fama internacional pedaleando durante una milla (1,609 km) a 101,8 km/h en una bicicleta, detrás de un tren de la compañía Long Island, sobre un camino construido para esa ocasión.

En 1912, Étienne Bunau-Varilla, de Francia, patentó un fuselado para bicicleta y ciclista, inspirado en la forma de los primeros dirigibles. Entre 1912 y 1933, las versiones de esta bicicleta y sus sucesoras establecieron los récords de velocidad en Europa. En 1933, el francés Marcel Berthet cubrió 49,98 kilómetros en una hora, montando un aparato fuselado llamado "Vélodyne". Su ritmo fue unos 5 km/h más veloz que cualquiera conseguido durante una hora en una bicicleta estándar.

Ese mismo año, el inventor francés Charles Mochet construyó una bicicleta de posición supina (el ciclista pedaleaba echado de espaldas) que, posteriormente, fuseló. Con el corredor profesional François Faure, este "Vélocar" estableció gran número de récords de velocidad entre 1933 y 1938. Mochet y Faure esperaban que los récords fueran reconocidos por la Unión Ciclista Internacional, el organismo mundial que rige las carreras ciclistas; pero no lo consiguieron.

En 1938, la Unión prohibía el uso de accesorios aerodinámicos y bicicletas de piloto acostado en las carreras: la norma continúa aún en vigor. Esta prohibición constituyó un serio obstáculo para el desarrollo de bicicletas de gran velocidad, y es una de las dos razones principales por las que la bicicleta ha permanecido casi sin cambios durante tanto tiempo. (La otra razón reside en la preferencia por el automóvil en los países desarrollados, donde la bicicleta ha perdido su anterior interés como medio de transporte.)

En su normativa, la Unión adscribía al capítulo de las "trampas" las mejoras en la aerodinámica y otros cambios tecnológicos. (Quizás haya sido una suerte que la Unión no existiera cuando el veterinario escocés-irlandés John Boyd Dunlop desarrolló las cubiertas para bicicletas de 1887, pues la gente podría estar montando bicicletas, e incluso automóviles, con sólidas ruedas de acero.) Sin embargo, hay que reconocer que la Unión ha empezado a relajar gradualmente sus restricciones sobre cambios en la aerodinámica, aunque la posición recostada del piloto sigue prohibida. Los trajes muy ceñidos y de una pieza se han hecho habituales en las carreras ciclistas internacionales desde 1976. Se ha permitido gorros aerodinámicos, secciones en forma de gota para los tubos del cuadro, palancas de freno carenadas y otros componentes mejorados aerodinámicamente. De hecho, los cambios tecnológicos en toda clase de vehículos de propulsión humana están hoy floreciendo a un ritmo desconocido desde el apogeo de la bicicleta en el siglo XIX.

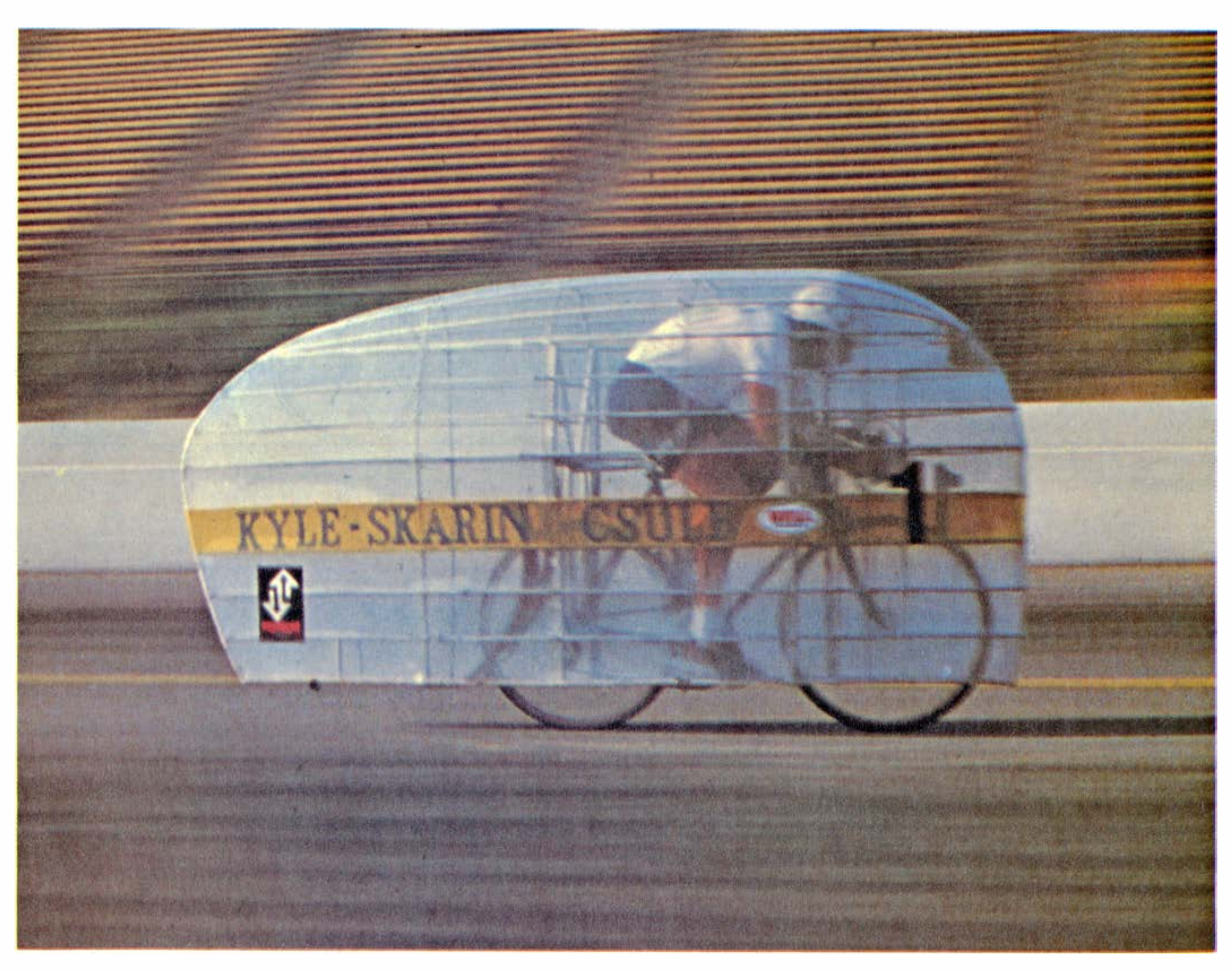
Este rápido cambio puede atribuirse, en parte, a una serie de hechos acontecidos en California. En 1973, uno de nosotros (Kyle) y Jack H. Lambie, consultor de aerodinámica, trabajando independientemente, construyeron y probaron las dos primeras bicicletas fuseladas en Estados Unidos. A diferencia de sus predecesores, Kyle y Lambie midieron la reducción en la resistencia conseguida con el fuselado. Para ello, realizaron numerosas pruebas a rueda libre, en las que se dejaba decelerar un vehículo no propulsado sobre una su-

perficie llana. En esas condiciones, la deceleración del vehículo es proporcional a la fuerza total de frenado que actúa sobre él; los instrumentos miden, bien la velocidad, bien la desaceleración. Kyle y Lambie, que publicaron sus resultados por separado, concluían que la fuerza de resistencia total en una bicicleta podría reducirse en más del 60 por ciento con un fuselado en forma de ala vertical que encerrase completamente a la bicicleta y al ciclista. (Hasta unos dos años después, ni Kyle ni Lambie sabrían que ya se habían construido en Europa vehículos similares.)

En 1974, en el puerto aeronaval de Los Alamitos, Ronald P. Skarin, ciclista olímpico de Estados Unidos, estableció cinco récords mundiales de velocidad montando la bicicleta fuselada de Kyle. A raíz de estos éxitos, Kyle y Lambie decidieron organizar una carrera para vehículos de propulsión humana sin restricciones. El 5 de abril de 1975, en Irwindale, California, 14 vehículos distintos compitieron en esta primera e histórica carrera. Muchos de ellos eran de piloto recostado, algunos con el corredor pedaleando en posición supina (cara arriba) y otros con el corredor en prono (cara abajo). Los había propulsados por manos y pies. El ganador, a 72,20 km/h, fue un tándem fuselado diseñado por Philip Norton, profesor de instituto en Near Claremont, California. Los ciclistas fueron el propio Norton y Christopher Deaton, corredor cualificado, pero no de talla mundial. (La bicicleta de carreras convencional más rápida sin ayuda ha alcanzado los 69,91 km/h, récord establecido en 1982 por el soviético Sergei Kopylov, ciclista de clase mundial.)

En 1976, enfrentados a la política de la Unión Ciclista Internacional contraria al fuselado, los participantes de esta carrera fundaron la Asociación Internacional de Vehículos de Propulsión Humana. Su propósito era reconocer las competiciones en las que los vehículos de propulsión humana no tuvieran limitaciones de diseño. A partir de entonces, en docenas de carreras realizadas en muchos países, las máquinas se han vuelto mucho más refinadas y las velocidades han aumentado sin parar. Cuatro vehículos han superado el límite de velocidad para automóviles en Estados Unidos de 88 km/h. (Recibiendo, todos ellos, una multa de honor, por exceso de velocidad, de la Patrulla de Carreteras de California.) Entre ellos se encuentra un cuadriciclo fuselado diseñado por Norton, de tercera generación.

Actualmente, el vehículo de propulsión humana más rápido del mundo es el "Vector Tandem", un biplaza de pilotos acostados, de elegante fuselaje. Fue construido por un equipo encabezado por Allan A. Voigt, inge-

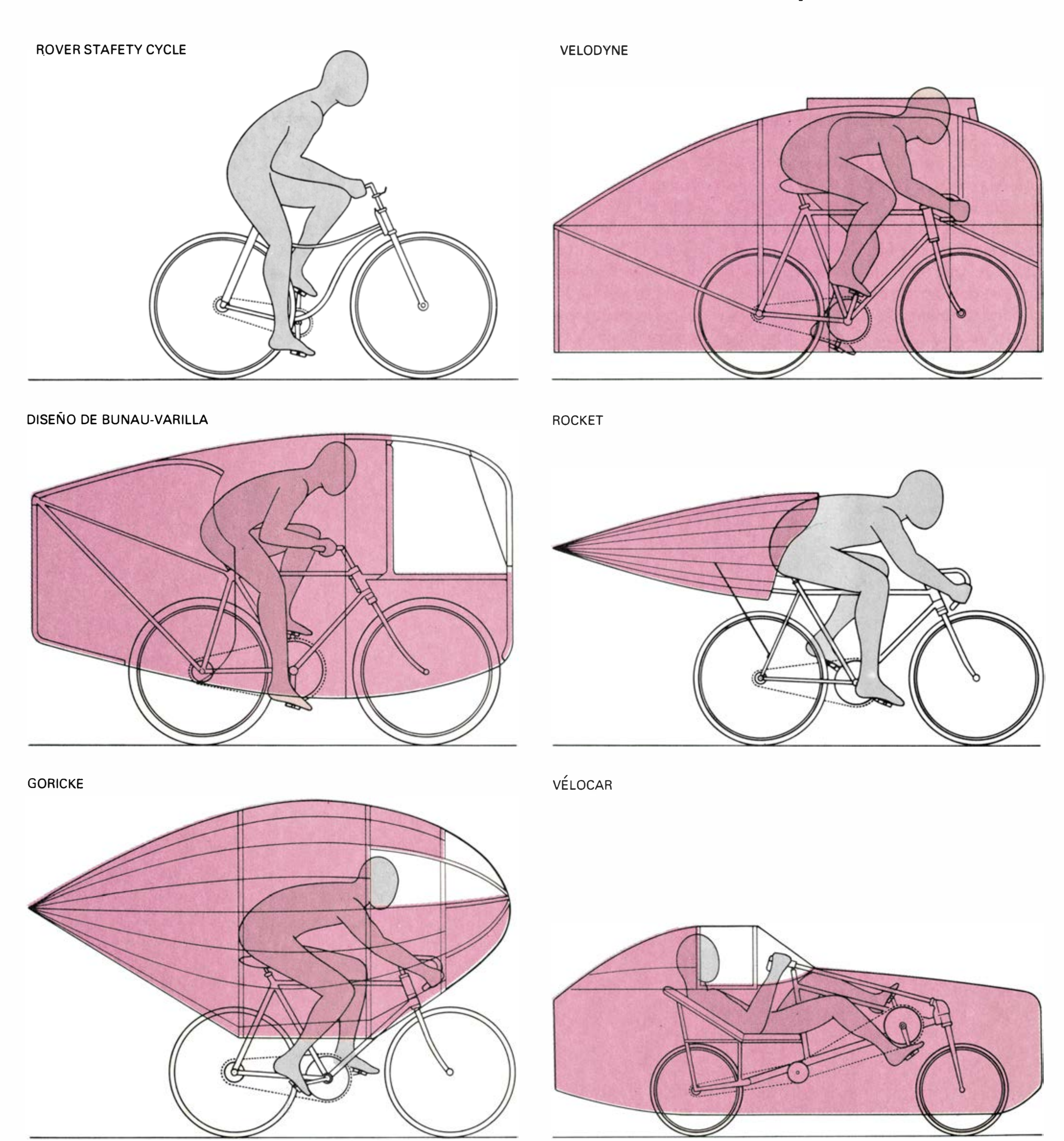


1. BICICLETA DE CARRERAS FUSELADA, diseñada por uno de los autores (Kyle) y montada por Ronald P. Skarin, ciclista olímpico estadounidense; se muestra aquí estableciendo el récord mundial de 51,3 kilómetros por hora durante una hora, con salida parada. La clave de esta marca estuvo en el fuselado aerodinámico, que redujo la resistencia aerodinámica del ciclista y la bicicleta. Skarin estableció el nuevo récord de velocidad en el año 1979, en la pista de carreras de Ontario (California). Si exceptuamos el fuselado, el vehículo venía a ser, fundamentalmente, una bicicleta de carreras convencional.

niero que, como presidente del Versatron Research, Inc., diseña principalmente servomotores aeroespaciales. (Los pedaleadores van en posición supina, mirando en direcciones opuestas.) En 1980, tras una toma de velocidad de alrededor de una milla (1,6 km), recorrió 200 metros a 101,24 km/h en la pista de carreras de Ontario, en California. A finales de ese año, el "Vector Tandem" consiguió una media de 81,3 km/h, durante 64 kilómetros, en la carretera de Stockton a Sacramento.

Estas extraordinarias velocidades son, casi por completo, producto del cuidado de la aerodinámica. Un ciclista que vaya a 32 km/h desplaza, normalmente, unos 450 kilogramos de aire por minuto. Cuando la máquina y el corredor no son aerodinámicos, dejan una estela considerable y pagan un alto coste en energía humana.

El rendimiento de una bicicleta se ve afectado por dos clases de resistencia

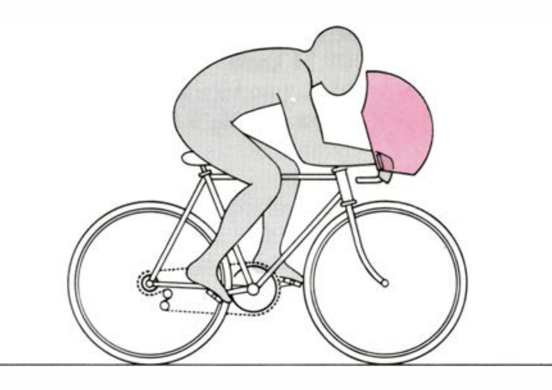


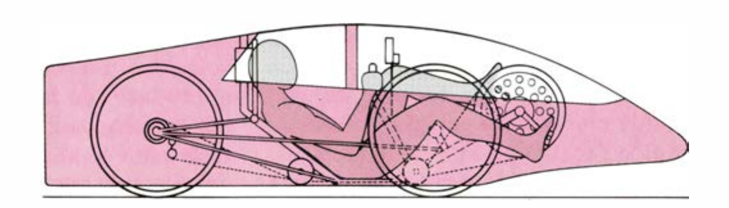
2. PRIMERAS MEJORAS incorporadas en vehículos terrestres de propulsión humana, con la introducción del "Rover Safety Cycle" en Inglaterra en 1884. En 1912 y 1913, el francés Étienne Bunau-Varilla obtenía las patentes para un fuselado; bicicletas parecidas establecieron muchos récords de velocidad. El "Goricke" se desarrolló en Alemania en 1914. El "Vélodyne" fue montado a 49,98 kilómetros por hora (nuevo récord) por el francés Marcel Berthet, en 1933. Del mismo año es el "Rocket", diseñado por Oscar Egg. Otro vehículo francés, el "Vélocar", estableció varios récords de velocidad entre 1933 y 1938. La mayoría de las ilustraciones de esta página están basadas en datos del Archivo Wolfgang Gronen de Binningen, Alemania.

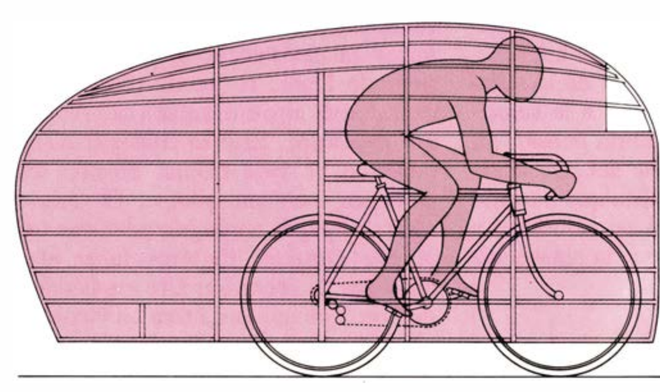
aerodinámica: resistencia de presión (o forma) y resistencia de fricción. La de presión se produce cuando el flujo de aire no logra adaptarse al contorno del objeto que se mueve. Esta separación varía la distribución de presión del aire sobre el objeto. Si esta separación tiene lugar hacia la parte de atrás del objeto, la presión del aire se hace allí menor que en la superficie delantera, produciendo resistencia.

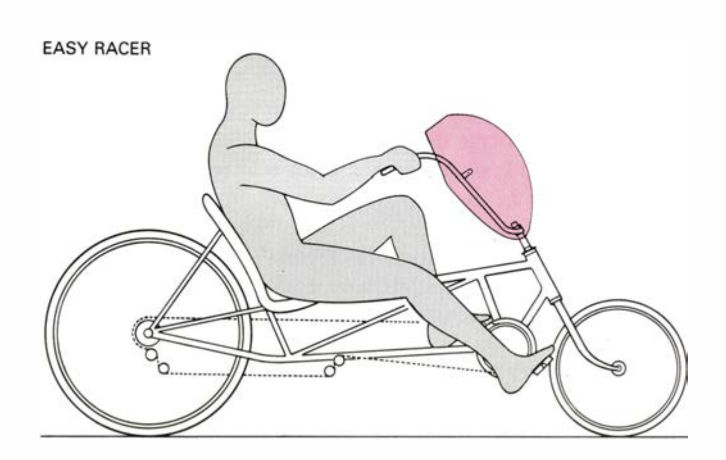
La resistencia de fricción se debe a la viscosidad del aire. Está producida por las fuerzas de cortadura que se crean en la capa límite, la capa de aire inmediata a la superficie del cuerpo.

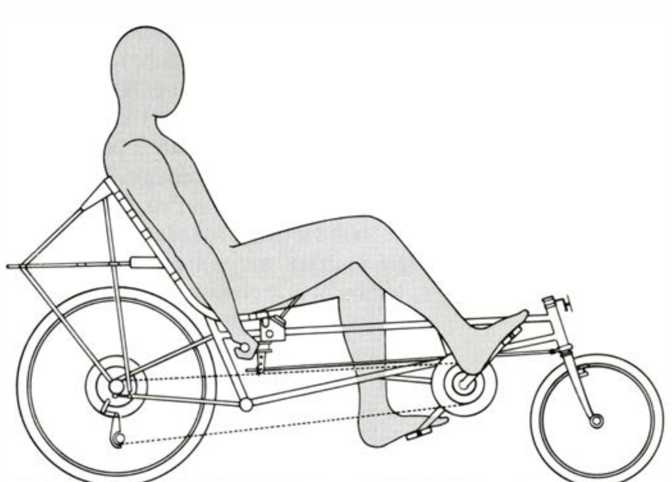
Las formas romas, tales como cilindros, esferas y otras formas habituales en una bicicleta, resultan ineficaces desde el punto de vista aerodinámico, por la sencilla razón de que el flujo de aire se separa de ellas. Se forman zonas de baja presión detrás de estos objetos, produciendo una resistencia de presión

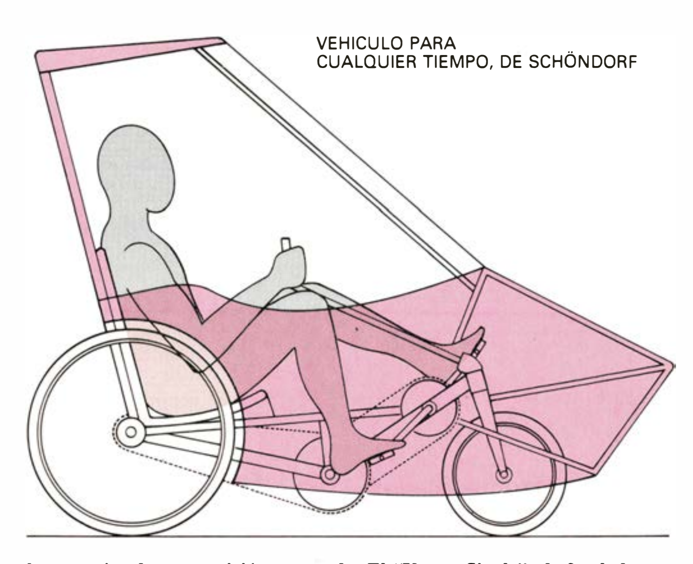












3. VEHICULOS DE PROPULSION HUMANA. Hacen uso intensivo del carenado para reducir la resistencia aerodinámica del conjunto vehículo-piloto. El más sencillo es el "Zzipper", con un fuselado parcial montado delante del ciclista. El "Kyle Streamliner" data de 1973. Un diseño, que sirve más para el turismo y el uso diario que para las carreras es el "Avatar 2000": aprovecha las ventajas de una posición recostada. El "Vector Single", de fuselado completo, alcanza casi 100 km/h, con una potencia del ciclista de un caballo de vapor. El "Easy Racer" es de piloto recostado, diseñado para turismo y uso diario; también ha competido. El último vehículo es uno de los de piloto recostado, diseñado por Paul Schöndorf, para la tercera edad y minusválidos.

centenares de veces mayor que la de fricción. Por el contrario, el aire fluye suavemente alrededor de las formas aerodinámicas, cerrándose detrás al pasar el objeto. La resistencia de presión se reduce bastante, pero la resistencia de fricción cobra mayor importancia.

Para una eficacia máxima, los vehículos deberían diseñarse minimizando la transferencia de energía irrecuperable al aire producida por estas dos clases de resistencia. En el nivel actual de tecnología, la resistencia aerodinámica absorbe del 40 al 50 por ciento de la energía del combustible consumido por un automóvil o un camión a 88 km/h. Puesto que la bicicleta tiene menor potencia, peso y resistencia de rodadura, y, además, una pobre aerodinámica, la resistencia absorberá, incluso, un porcentaje mayor de la energía consumida a velocidades superiores a 16 km/h.

El coeficiente de resistencia designa el rendimiento aerodinámico de un perfil. Un perfil ineficiente, una esfera, tendrá un coeficiente de resistencia de, digamos, 1,3, mientras que una forma aerodinámica, la de una gota, tendrá uno menor de 0,1. Por tanto, un objeto con un perfil en forma de gota se moverá con una pérdida de energía inferior a una décima de la de otro objeto que posea forma cilíndrica.

En los vehículos de transporte terrestre, la resistencia aerodinámica es, casi, directamente proporcional al producto del área frontal por el coeficiente de resistencia. Por convenio, llamaremos área frontal efectiva a este producto. Para saber cuál de dos vehículos tiene menor resistencia aerodinámica, no basta comparar sus coeficientes; hay que tener también en cuenta el tamaño de la máquina. Esto se consigue con el concepto de área frontal efectiva. Una bicicleta corriente, con su ciclista, tendrá un área frontal eficaz de entre 0,3 y 0,6 metros cuadrados, mientras que un vehículo terrestre de propulsión humana fuselado puede no llegar a 0,045 metros cuadrados.

La fuerza de resistencia aerodinámica aumenta con el cuadrado de la velocidad. La potencia es proporcional al producto de la fuerza de resistencia por la velocidad; así, la potencia necesaria para conducir un objeto a través del aire aumenta con el cubo de la velocidad. Por tanto, un pequeño incremento de velocidad requiere un gran aumento de potencia. Un ciclista que doble repentinamente su potencia cuando va a 32 km/h aumentará su velocidad apenas hasta unos 42 km/h.

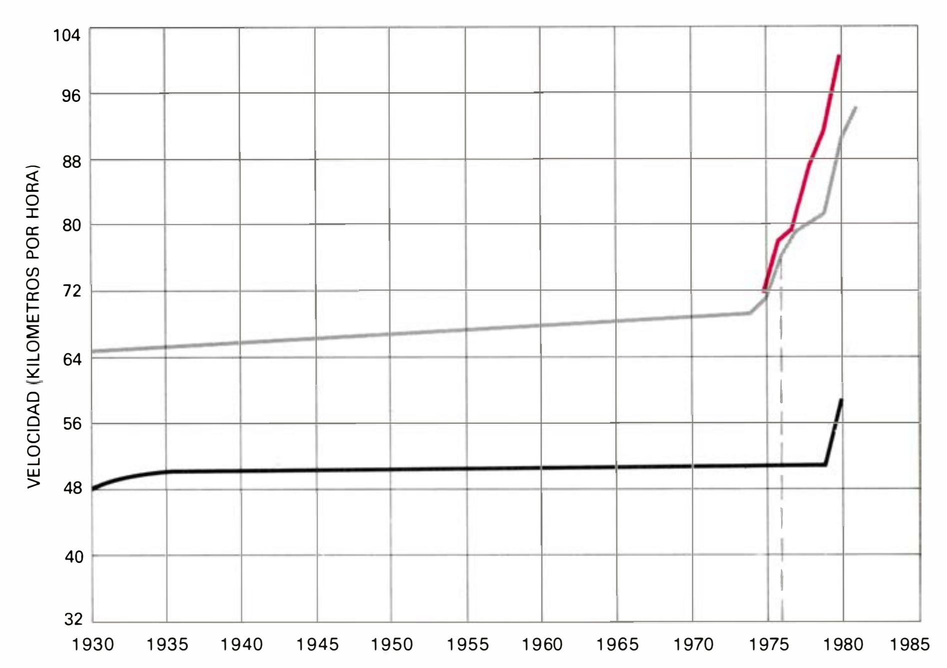
En cambio, la reducción de la resistencia aerodinámica afecta a la velocidad menos de lo que cabría pensar. Si, a 32 km/h, la resistencia del aire se reduce a la mitad, un ciclista que no varíe su potencia aumentará su velocidad a 39 km/h. La razón es que la resistencia de rodadura permanece constante. Si pudiera despreciarse esta resistencia, doblando la potencia o reduciendo a la mitad el área frontal efectiva, conseguiríamos que la velocidad alcanzara otra vez los 42 km/h.

En resumen, las velocidades altas exigen un rendimiento aerodinámico extremadamente alto. El "Vector Tandem", con una potencia ligeramente mayor a un caballo de vapor por cada ciclista, alcanzó una velocidad de 101,24 km/h. Para conseguir esta velocidad, una bicicleta convencional necesitaría más de 6 CV, nivel de potencia claramente imposible para un ser humano.

Los diseñadores y los ciclistas pueden reducir de tres formas principales la resistencia aerodinámica en los vehículos de propulsión humana. En primer lugar, pueden reducir la cantidad de energía malgastada en la interacción del vehículo con el aire. Esto se hace fuselando (perfilando el frente y el dorso de los objetos romos para minimizar la resistencia de presión) y suavizando las superficies rugosas para minimizar la resistencia de fricción. En segundo lugar, se puede disminuir la cantidad de aire enfrentada cada segundo de avance. Esto se consigue disminuyendo el área frontal efectiva del conjunto vehículo-piloto. El mismo efecto se puede conseguir pedaleando a grandes altitudes. En tercer lugar, el ciclista puede encontrar aire moviéndose de forma tal que produzca un viento de cola. El ejemplo más claro lo tiene cuando se deja arrastrar, esto es, cuando corre muy cerca de la estela de otro vehículo.

A grandes altitudes, la atmósfera es menos densa y los ciclistas se enfrentan con menos aire. En la ciudad de México (de 2260 metros de altitud, donde la densidad del aire es sólo el 80 por ciento de la del nivel del mar) los récords ciclistas son de un 3 a un 5 por ciento más rápidos que los conseguidos a altitudes inferiores. En La Paz, Bolivia (altitud 3600 metros), los récords de velocidad podrían mejorarse, teóricamente, en un 14 por ciento. En la Luna, sin atmósfera y con un sexto de atracción de la gravedad, un ciclista convenientemente equipado podría correr a 383 km/h con la muy modesta potencia de 0,1 CV.

Al analizar el 80 por ciento de la potencia generada por un ciclista que, al moverse en suelo llano a 29 km/h, se gasta en vencer la resistencia del aire, encontramos que alrededor de un 70 por ciento de este consumo se debe a la



4. RECORDS DE VELOCIDAD de los vehículos terrestres de propulsión humana, en rápido crecimiento desde la formación, en 1976, de la Asociación Internacional de Vehículos de Propulsión Humana, que no pone limitaciones de diseño. El año de su fundación está indicado con la línea a trazos. Hasta entonces, las normas de la Unión Ciclista Internacional, que prohibía los vehículos fuselados para las competiciones ciclistas reconocidas habían mantenido los récords de velocidad prácticamente inalterados. Las curvas representan los récords para vehículos múltiples, sobre 200 m, con toma de velocidad (*color*), corredores en las mismas condiciones (*gris*) y corredores pedaleando una hora con el máximo esfuerzo (*negro*).

resistencia del aire al ciclista y el 30 por ciento a la de la bicicleta. Esto nos lleva a la conclusión de que, para mejorar las actuaciones de una bicicleta estándar, hay que perfeccionar primero la aerodinámica del ciclista.

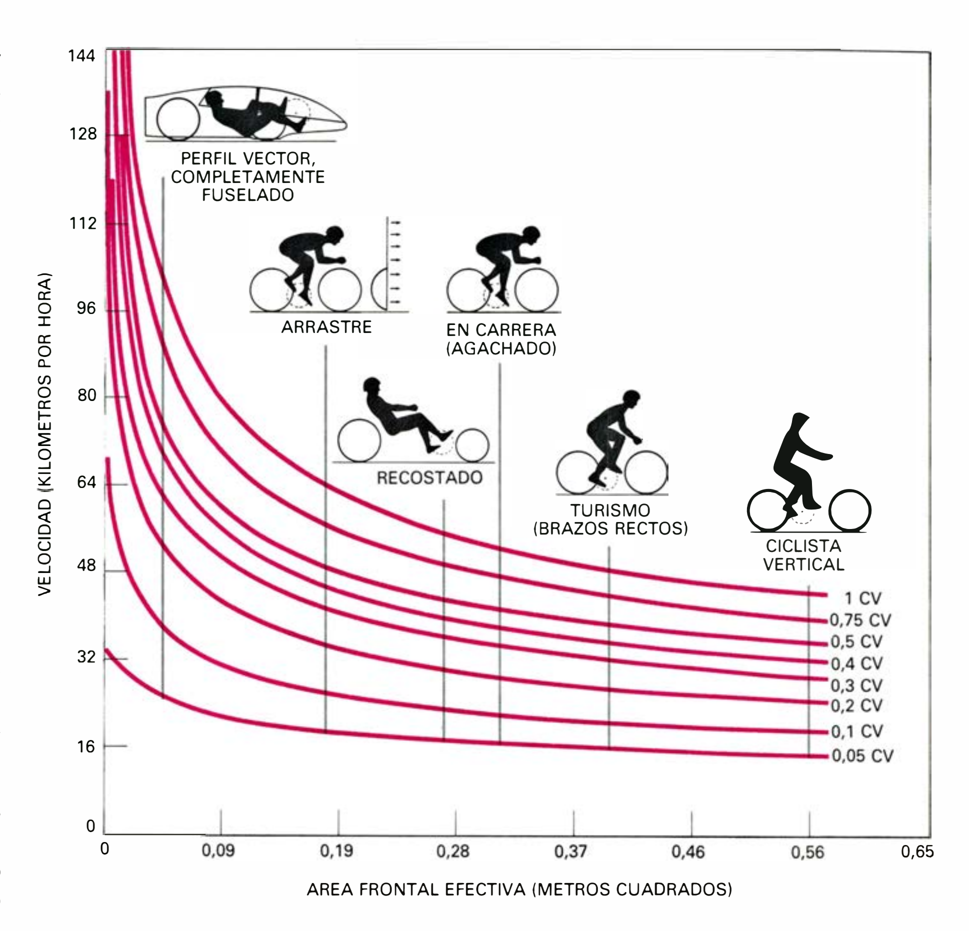
Para los ciclistas de competición, las limitaciones de la Unión Ciclista Internacional dejan poco espacio para mejoras, después de las ya conseguidas con la adopción de la posición en cuclillas, el casco aerodinámico, el traje ceñido y el perfilado de los componentes de la bicicleta. Voigt ha calculado que, incluso con una bicicleta "perfecta" (sin resistencia aerodinámica a cualquier velocidad y cubiertas sin resistencia de rodadura), la sola resistencia aerodinámica del piloto obstaculizaría gravemente la mejora de los rendimientos. Según Voigt, un ciclista en cuclillas sobre una bicicleta de carreras convencional podría alcanzar una velocidad máxima de unos 55 km/h con una potencia de 1 CV. En una bicicleta perfecta, el mismo corredor, con el mismo esfuerzo, podría conseguir 61 km/h.

Para los millones de ciclistas que no participan en las competiciones, y sólo quieren sacarle más rendimiento a su bicicleta, caben algunas mejoras aerodinámicas. Estas pueden clasificarse según su coste, empezando por la más barata: un carenado parcial como el del "Zzipper", desarrollado y fabricado por Glen Brown, de Santa Cruz, California. Es una placa aerodinámica, pequeña y transparente, montada delante del ciclista. Por unos 60 dólares (9600 pesetas), un ciclista puede disminuir la resistencia aerodinámica en un 20 por ciento, consiguiendo un incremento de velocidad de unos 4 km/h, para un CV de potencia.

Otra forma eficaz de reducir la resistencia aerodinámica consiste en usar una bicicleta de piloto recostado. (La máquina puede costar algunos cientos de dólares más que una bicicleta de aficionado. Los pioneros en este campo son Gardner Martin, de Freedom, California, diseñador del "Easy Racer", y David Gordon Wilson, del Instituto de Tecnología de Massachusetts, diseñador del "Avatar 2000". La resistencia del viento disminuye de un 15 a un 20 por ciento, a causa de la menor área frontal que presenta el ciclista tumbado, consiguiendo aproximadamente el mismo incremento de velocidad que con el carenado del "Zzipper".

Sin embargo, la bicicleta de piloto recostado ofrece otras ventajas. Es más cómoda que una convencional. Mucho más segura en accidentes que no sean de colisión con automóviles, ya que el ciclista se halla más próximo al suelo (haciendo menos grave la caída) y los pies están por delante (haciendo menos probable una lesión en la cabeza en caso de caída). Tiene el problema de que es difícil de ver en la carretera y, por ello, resulta quizá más vulnerable a los automóviles; problema que puede mitigarse en parte colocando en el vehículo un mástil delgado con un banderín.

En el extremo más caro de la escala se encuentra una bicicleta con fuselado completo. El "Vector Single", versión individual del "Vector Tandem", es el mejor ejemplo de fuselado completo y cerrado de vehículo a pedal. (Se trata de la máquina representada en la portada de este número de *Investigación y Ciencia*.) Según Voigt, el vehículo es teóricamente capaz de alcanzar 99,3 km/h con 1 CV de potencia, 45,4 km/h más de lo conseguido con una bicicleta de carreras estándar. Un "Vector Single" viene a costar como una bicicleta de carreras de primera clase.



5. EFECTO DEL FUSELADO, mejorando los rendimientos de los vehículos de propulsión humana para cualquier nivel de potencia. El de ciclista vertical es el vehículo menos aerodinámico; el perfil "Vector", el más aerodinámico. El arrastre consiste en seguir muy de cerca a otro vehículo, aquí una bicicleta. Un buen atleta puede desarrollar 1 CV durante unos 30 segundos; un ciudadano normal lo hace a lo largo de 12 segundos. Ambos pueden mantener 0,4 y 0,1 CV, respectivamente, durante unas ocho horas. Se llama área frontal efectiva al producto del coeficiente de resistencia por la proyección del área frontal.

Un vehículo totalmente fuselado mantiene sus ventajas sobre una bicicleta convencional en subidas y bajadas. Aunque el "Vector Single" pesa unos 36 kilogramos, frente a los 11 de una bicicleta estándar, puede subir pequeñas colinas tan ligero o más que una bicicleta. Con una potencia de 0,4 CV, una bicicleta remonta pendientes del 2,5 por ciento a unos 26 km/h y del 6 por ciento a unos 18 km/h. Con la misma potencia, el Vector puede subir las dos clases de pendiente a 33 y 18 km/h respectivamente.

En bajada, la diferencia entre una y otra máquina es considerable. Una bicicleta puede descender por una pendiente del 2,5 por ciento a 47,25 km/h; el Vector a 86,9. En pendientes del 6 por ciento, la bicicleta alcanza una velocidad de 62,48 km/h mientras el Vector supera los 160. Tales velocidades potenciales implican que, caso de generalizarse los vehículos de propulsión humana fuselados, deberá ponerse especial cuidado en el diseño de los frenos, la suspensión y la estabilidad del vehículo.

	DESCRIPCION		FUERZAS A 32 KM/H (EN KILOS)	DATOS AERODINAMICOS		
				COEFICIENTE DE RESISTENCIA	AREA FRONTAL (EN METROS CUADRADOS)	AREA FRONTAL EFECTIVA (M^2)
BICICLETAS CONVENCIONALES	BMX (JOVEN CORRE- DOR DE CROS)	BICICLETA: 13,5 kg PILOTO: 54 kg NEUMATICOS DE TACOS DIAMETRO 51 cm, 2,8 kg/cm^2	2,48 0,94	1,1	0,46	0,50
	USUARIO EUROPEO VERTICAL	BICICLETA: 11 kg PILOTO: 72,5 kg NEUMATICOS, DIAMETRO 68 cm, 2,8 kg/cm^2	2,76 0,54	1,1	0,51	0,56
	TURISMO (BRAZOS RECTOS)	BICICLETA: 11 kg PILOTO: 72,5 kg NEUMATICOS NORMALES DIAMETRO 68 cm, 6,3 kg/cm^2	1,98 0,37	1	0,40	0,40
	CARRERA (EN CUCLILLAS)	BICICLETA: 9 kg PILOTO: 72,5 kg NEUMATICOS DE CARRERA DIAMETRO 68 cm, 7,4 kg/cm^2	1,57 0,24	0,88	0,36	0,32
MODELOS MEJORADOS	COMPONENTES AERODINAMICOS (EN CUCLILLAS)	BICICLETA: 9 kg PILOTO: 72,5 kg NEUMATICOS DE CARRERA DIAMETRO 68 cm, 7,4 kg/cm^2	1,47 0,24	0,83	0,36	0,30
	CARENADO PARCIAL (ZZIPPER, EN CUCLILLAS)	BICICLETA: 9,5 kg PILOTO: 72,5 NEUMATICOS DE CARRERA DIAMETRO 68 cm, 7,4 kg/cm^2	1,34 0,24	0,70	0,38	0,27
	RECOSTADO (EASY RACER)	BICICLETA: 12 kg. PILOTO: 72,5 kg NEUMATICOS NORMALES DELANTERA: 51 cm TRASERA: 68 cm, 6,3 kg/cm^2	1,34 0,42	0,77	0,35	0,27
	TANDEM	BICICLETA: 19 kg PILOTO: 72,5 kg, DOS NEUMATICOS NORMALES 68 cm DIAMETRO, 6,3 kg/cm^2	2,39 1,20 0,73 0,36	1	0,48	0,48
	ARRASTRE (SIGUIENDO MUY DE CERCA A OTRA BICICLETA)	BICICLETA: 9 kg PILOTO: 72,5 kg NEUMATICOS DE CARRERA DIAMETRO 68 cm, 7,4 kg/cm^2	0,87 0,24	0,50	0,36	0,18
POSEEDORES DE RECORDS	BLUE BELL (DOS RUEDAS, UN PILOTO)	BICICLETA: 18 kg. PILOTO: 72,5 kg NEUMATICOS DE CARRERA, DELANTERA: 51 cm TRASERA: 68 cm, 7,4 kg/cm^2	0,27 0,36	0,12	0,46	0,06
	KYLE (DOS RUEDAS, DOS PILOTOS)	BICICLETA: 24 kg PILOTO: 72,5 kg, DOS NEUMATICOS DE CARRERA, 7,4 kg/cm^2	0,65 0,32 0,50 0,25	0,2	0,65	0,13
	VECTOR SINGLE (TRES RUEDAS)	BICICLETA: 31 kg. PILOTO: 72,5 kg NEUMATICOS DE CARRERA DELANTERA: 61 cm TRASERA: 68 cm	0,23 0,46	0,11	0,424	0,046
	VECTOR TANDEM (TRES RUEDAS)	BICICLETA: 34 kg PILOTO: 72,5 kg, DOS NEUMATICOS DE CARRERA, DIAMETRO 61 cm	0,28 0,14 0,80 0,40	0,13	0,44	0,056
LIMITES TEORICOS	BICICLETA PERFECTA	SIN RESISTENCIA DE RODADURA, SIN RESISTENCIA SOBRE LA BICICLETA	1,38 0	0,8	0,35	0,28
	CORREDOR SIN RESISTENCIA	LA RESISTENCIA DE RODADURA INCLUYE EL PESO DEL CORREDOR	0,60 0,36	1,1	0,11	0,12
	RECOSTADO PERFECTO	RESISTENCIA SOBRE EL PILOTO SOLAMENTE	0,32 0	0,6	0,11	0,06
	BICICLETA PRONO PERFECTA	RESISTENCIA SOBRE UN PILOTO PEQUEÑO PERO FUERTE	0,23 0	0,6	0,074	0,046
	FUSELADO PRONO PERFECTO		0,03 0	0,05	0,13	0,006
	TRAS MOTO	BICICLETA: 19 kg PILOTO: 72,5 kg NEUMATICOS DE MOTO DE CARRERAS, 4,9 kg/cm^2	0 0,54			VARIA CON LA VELO- CIDAD
	BICICLETA LUNAR	BICICLETA: 11 kg PILOTO: 72,5 kg TRAJE LUNAR: 7 kg	0 0,07			0

RESISTENCIA DE RODADURA	SUELO LLANO, SIN VIENTO			EFECTO DE LAS PENDIENTES	
	POTENCIA NECESARIA	TRAYECTO DIARIO	VELOCIDAD MAXIMA	VELOCIDAD ESTACIONARIA EN SUBIDA	VELOCIDAD ESTACIONARIA A RUEDA LIBRE
0,014	146	16,2	44,7	19,6	31,86
0,006	140	18,2	44,4	17,5	38,6
0,0045	100	21,1	50	19,6	44,6
0,003	77	23,6	54,5	20,9	50,2
0,003	73	24,1	55,7	20,9	51,8
0,003	67	24,8	57,4	21,1	54,5
0,005	75	23,2	56,6	20,1	54,2
0,0045	66	24,5	58,9	20,9	56,6
0,003	47	28,2	66	21,9	67,1
0,004	27	36,2	94,3	20,8	124,5
0,003	24	37,5	91,1	22,5	112,5
0,0045	29	35,1	98,5	18,2	145
0,0045	23	41,2	116,6	20,9	174,4
0	59	26,9	57,8	21,6	55,8
0,0045	41	29,6	73,7	21,4	80,9
0	14	43,6	93,8	27	107,6
0	10	48,9	105,1	37,3	105,1
0	1	93,8	202,6	41,2	280,8
0,006	23	47,3	473	20,3	?
0,0045	3	382,1	3821	126,2	?

Dado que la fuerza de resistencia aerodinámica es proporcional al cuadrado de la velocidad relativa, el viento en contra, el de cola e incluso el viento cruzado pueden variar drásticamente tanto la resistencia aerodinámica como las necesidades de potencia. Por ejemplo, un ciclista que vaya a 29 km/h con el aire tranquilo deberá aumentar su esfuerzo en un ciento por ciento para mantener esa velocidad contra un viento de 16 km/h. Cuando un ciclista se enfrenta a un viento de cara, generalmente reduce la marcha y trata de mantener su ritmo de pedaleo y su esfuerzo habitual cambiando el desarrollo (la marcha). Esta es una de las razones por las que las bicicletas con desarrollos múltiples son interesantes, incluso en países llanos.

El viento de cola hace que, con su esfuerzo habitual, el ciclista vaya más rápido. En general, el aire en movimiento acelerará o frenará la bicicleta en, aproximadamente, la mitad de la velocidad del viento. Cuando un ciclista va en la estela de otro, las necesidades de potencia del que es arrastrado se reducen en un 30 por ciento. El ciclista en cabeza crea un viento de cola artificial.

Cuanto más de cerca siga la bicicleta de atrás a la primera, más pronunciado será el efecto de arrastre. Se puede considerar al corredor trasero de un tándem como arrastrado desde muy cerca; de hecho, los ciclistas de tándem utilizan un 20 por ciento menos de potencia por corredor que dos ciclistas separados.

Si los ciclistas van en fila, turnándose en la posición de cabeza, el grupo entero irá mucho más deprisa que un corredor aislado. En una carrera de persecución de 4000 metros, un equipo de cuatro corredores puede ir unos 6 km/h más rápido que un ciclista aislado. Normalmente, un grupo de ciclistas aficionados de igual potencia puede ir de 1,5 a 5 km/h más rápido que cualquiera de ellos solo. Cuanto mayor sea el grupo, más rápido debería ser capaz

6. RENDIMIENTOS de los vehículos de propulsión humana, resumidos en el esquema adjunto. Los números de la columna de fuerzas representan la resistencia del aire y la de rodadura para cada vehículo, respectivamente. Las cinco columnas a la derecha simbolizan, sucesivamente, la potencia necesaria, a 32 km/h, en porcentaje al rendimiento del ciclista aficionado; la velocidad media en un habitual trayecto diario, en km/h, con 0,1 CV de potencia; la velocidad máxima con 0,1 CV; la velocidad estacionaria en km/h subiendo una pendiente del 5 por ciento con 0,4 CV de potencia y la velocidad a rueda libre bajando esa pendiente.

de correr (hasta, digamos, una docena de ciclistas).

Los vientos artificiales creados por el tráfico de automóviles pueden aumentar la velocidad de un ciclista entre 1,5 y 5 km/h durante períodos de unos siete segundos. Cuanto mayor sea el vehículo que pasa, más se nota su efecto. Una corriente de tráfico estacionaria puede permitir a un ciclista mantener una velocidad de 5 a 10 km/h por encima de la que le sería posible, de otro modo, para un gasto dado de energía.

Cuando un ciclista marcha justo en la estela de un vehículo de motor, puede conseguir velocidades bastante notables. A esta modalidad se la llama "tras moto". El 25 de agosto de 1973, Allan V. Abbott, médico de California, consiguió un récord de 223,126 km/h durante una milla (1,609 km) tras moto, en Bonneville Salt Flats, en Utah. John Howard, ciclista olímpico de Estados Unidos, se ha propuesto romper el récord de Abbott y alcanzar velocidades tras moto superiores a 240 km/h.

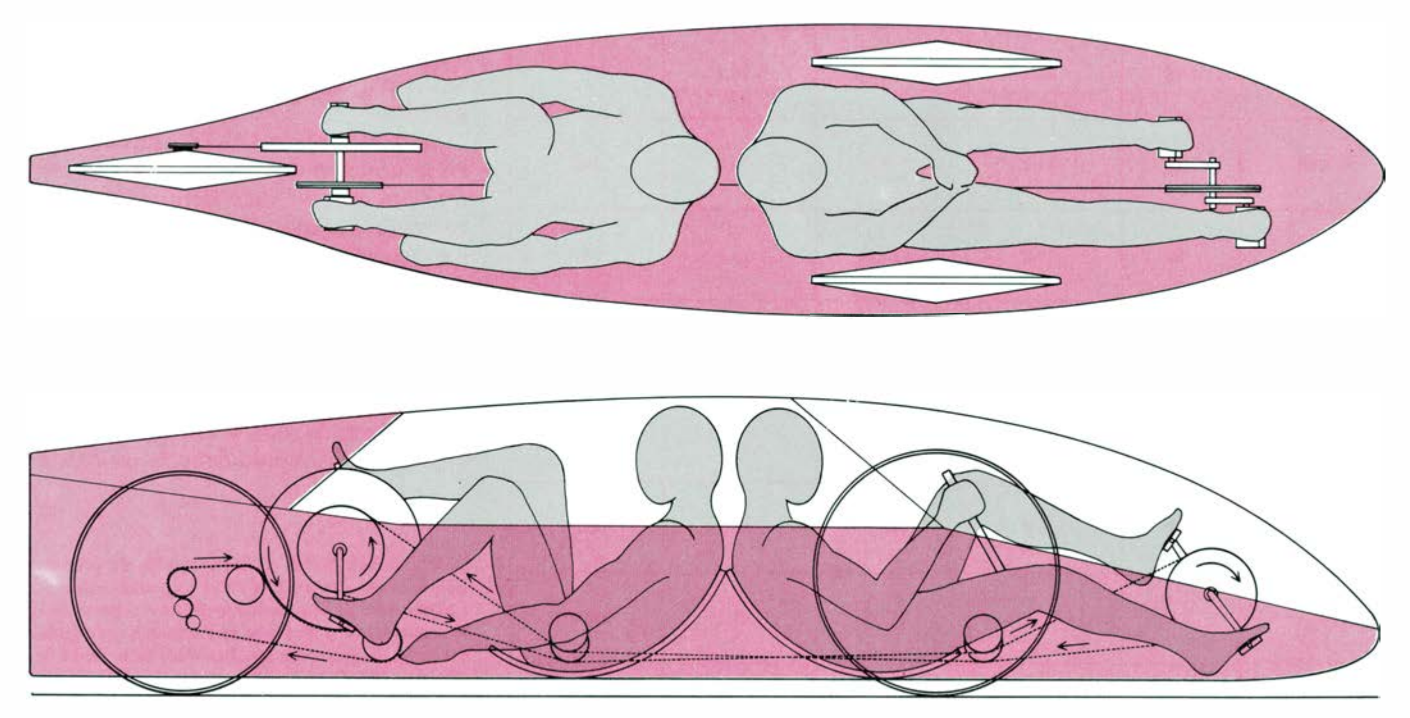
Aunque los resultados que hemos descrito son por sí mismos significativos, cabe preguntarse si tendrán alguna aplicación práctica, más allá de su efecto sobre los récords de velocidad. Parece improbable que estos trabajos ofrezcan una utilidad inmediata para gran parte de los ciclistas del mundo. En muchos países en vías de desarrollo, donde la bicicleta constituye el principal medio de transporte, la mayoría de los conductores van a unos 11 km/h, a menudo con una carga considerable; pero la resistencia aerodinámica sólo se hace más importante que otros obstáculos al movimiento de la bicicleta a velocidades por encima de los 16 km/h. Pero también aquí el conocimiento de la aerodinámica aporta su contribución: sin él, los diseñadores no sabrían por qué deben despreciar la aerodinámica en los vehículos de propulsión humana lentos.

Para las bicicletas, destinadas a un progreso lento pero seguro, tiene sentido disminuir la resistencia de rodadura mejorando las cubiertas y pavimentando las carreteras. Los diseñadores deberían reducir el peso de la máquina para facilitar la subida de pendientes. La reciente introducción de las "bicicletas de montaña" en los Estados Unidos constituye un avance en la fabricación de tipos ligeros, pero bastante resistentes, para los caminos accidentados o sin pavimentar.

Los conocimientos adquiridos a través de las últimas investigaciones sobre la aerodinámica de los vehículos de propulsión humana pueden aplicarse directamente de varias formas. Aunque es posible que la bicicleta tradicional persista dominando durante muchos años, por su propia aceptación, bajo coste, simplicidad y fiabilidad mecánica, ofrece muchos campos a la innovación. Por ejemplo, un sencillo fuselado frontal, ligero y barato, mejoraría las prestaciones de la bicicleta estándar. Las bicicletas de piloto recostado pueden extenderse entre los usuarios habituales y los turistas, debido a su eficacia y comodidad.

Una aplicación de la tecnología sería añadir un motor pequeño y ligero, de no muchos caballos de potencia, a la de piloto recostado. El motor serviría, principalmente, de ayuda para arrancar y subir pendientes. Si se le añade también todo el fuselado que permitan las necesidades de ventilación y estabilidad, la máquina se convertiría en un auténtico ciclomotor. (Los que se venden ahora con ese nombre no son realmente vehículos a pedal con motor, sino motocicletas de baja potencia.)

Las investigaciones han sugerido a los inventores el desarrollo de vehículos de propulsión humana específicos. Paul Schöndorf, profesor de ingeniería en Colonia, ha construido una serie de triciclos de piloto recostado, de fácil pedaleo en cualquier tiempo, para la tercera edad y los minusválidos. Vehículos de este tipo harían buen servicio en las comunidades de jubilados. Douglas Schwandt, del Centro de Investigación y Desarrollo de Ingeniería de Rehabilitación de Veteranos de la Administración, en Palo Alto, California,



7. "VECTOR TANDEM", presentado aquí en planta y alzado. Es el compañero del "Vector Single", que se ofrece en la portada de este número de INVESTIGACIÓN Y CIENCIA. El Tándem, con una potencia algo mayor de 1 CV para cada uno de los dos pilotos, situados espalda contra espalda, estableció la velocidad récord de 101,24 km/h, sobre 200 metros, en 1980 (los corredores tomaron velocidad durante más de una milla, 1,609 km). A finales de ese año, en una autopista interestatal de California, el "Vector Tandem" consiguió una velocidad media de 81,2 km/h en un trayecto de 64 kilómetros.

ha ideado triciclos de manivela y bicicletas para parapléjicos. William Warner, un parapléjico que alcanzó el récord de vehículos de propulsión a mano en competiciones patrocinadas por la Asociación Internacional de Vehículos de Propulsión Humana, piensa que una persona incapacitada puede impulsar estos vehículos a mucha mayor velocidad que una silla de ruedas convencional, ganando así un nuevo sentido de libertad y movilidad. (El actual récord de 40,37 km/h fue establecido, en 1981, por Ascher Williams, del centro de rehabilitación de Palo Alto.)

En principio, un vehículo de propulsión humana con fuselado completamente cerrado podría ser bastante útil para el transporte. El ciclista se desplazaría a velocidades de 32 a 48 km/h con cualquier clase de tiempo. Sin embargo, tal y como se diseñan ahora esos vehículos, no servirían para la circulación rodada. Carecen de ventilación, visibilidad y maniobrabilidad adecuadas, así como medios de seguridad imprescindibles (luces y limpiaparabrisas). En la mayoría de ellos no es fácil entrar ni salir.

Producir un vehículo práctico de esta clase requeriría una inversión y un esfuerzo en ingeniería comparables a los realizados en la producción de nuevos automóviles. Aun así, el vehículo a pedales no sería seguro en un tráfico con gran número de vehículos motorizados. Podríamos concluir que el vehículo de propulsión humana completamente cerrado no será una forma práctica de transporte hasta que la falta de combustible eche a la mayoría de los vehículos motorizados de la carretera o hasta que se construyan calzadas especiales para las máquinas a pedales.

Mucho más probable es el empleo de gran parte de la tecnología que hemos descrito en el desarrollo de automóviles más ligeros y de menor consumo. Uno de los autores (Malewicki) ha construido ya un vehículo de ese tipo, una máquina individual de 104 kilogramos de peso, que tiene récords de economía en combustible, a la velocidad en pista libre de 88 km/h, con motor de gasolina (66,45 kilómetros por litro) y diesel (66,07 km/l). El récord en diesel fue establecido en un viaje de Los Angeles a Las Vegas, con una velocidad media de 90,6 km/h. La moda hacia estos vehículos podría ayudar a prolongar los recursos de combustible posponiendo, irónicamente, el momento en que los vehículos de propulsión humana ocuparán el lugar que les corresponde.

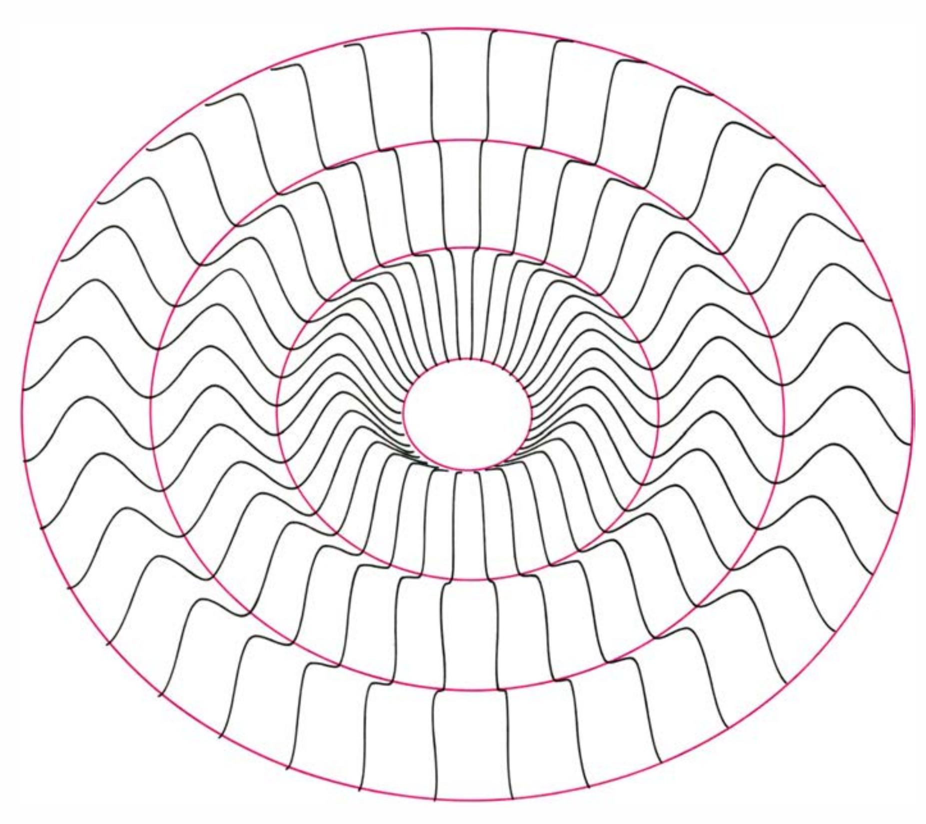
Interpretación de las ilusiones ópticas

Aparentemente, el sistema de la visión organiza imágenes retinianas ambiguas según reglas de inferencia que se aprovechan de ciertas regularidades del mundo exterior

Donald D. Hoffman

La visión es un proceso inferencial. Lo que se ve cuando miramos en derredor depende no sólo de las cosas visibles, sino también de cómo organiza nuestro sistema visual las imágenes que impresionan las retinas. Una curiosa demostración de este aspecto de la percepción nos la proporciona la aparente superficie que se forma al hacer girar una onda cosenoide en torno a un eje vertical y mirársela oblicuamente [*véase la figura 1*]. Al observar por primera vez la figura nos parece organizada en una serie de anillos concéntricos en relieve, con los límites entre los anillos marcados aproximadamente por las líneas circulares de color. Pero si giramos la página de modo que las letras queden en posición invertida, de abajo arriba, entonces la organización de la figura cambia: las líneas de color no caen ahora en el fondo de los pliegues entre anillos, sino que parecen seguir las crestas o salientes de los mismos. (Compruébelo.) Es obvio que el sistema visual hace algo más que transmitir pasivamente señales al cerebro: toma parte activa en organizarlas e interpretarlas.

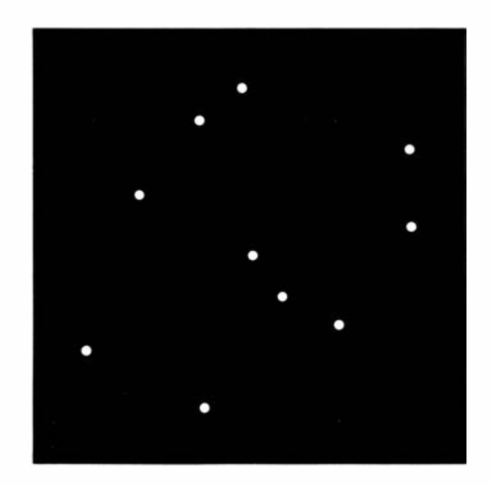
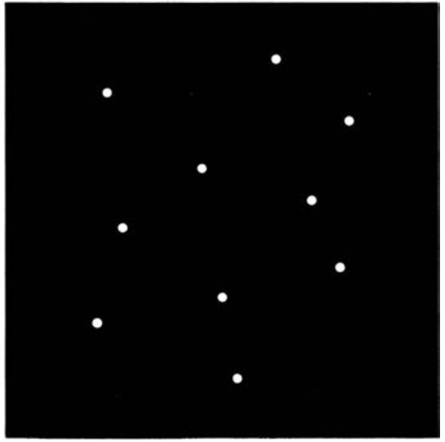
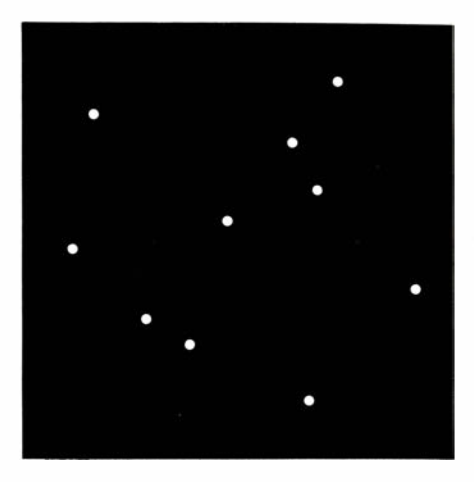
Este hallazgo plantea tres cuestiones. En primer lugar, ¿por qué el sistema visual necesita organizar e interpretar las imágenes formadas en las retinas? En segundo lugar, ¿cómo en tal proceso sigue siendo fiel al mundo real? Y, en tercer lugar, ¿qué reglas de inferencia cumple? Para responder a estas preguntas hay que examinar más detenidamente las figuras.



1. SUPERFICIE AMBIGUA obtenida mediante la rotación de una onda cosenoidal en torno a un eje vertical. Inicialmente la superficie parece estar organizada en forma de anillos en relieve y concéntricos, con las líneas rojas circulares en los valles. Pero cuando se invierte la página, haciéndola girar de arriba abajo, la organización parece cambiar: los círculos rojos parecen caer entonces en la cresta de los anillos.

Una razón de que el sistema visual organice e interprete las imágenes retinianas es, simplemente, que con cualquier imagen retiniana concreta son compatibles muchas posibles configuraciones del mundo real. Dicho de otro modo, las imágenes retinianas necesitan organización e interpretación porque, fundamentalmente, son ambiguas. Ambigüedad que nace, en parte, de la tridimensionalidad del mundo, en tanto que cada imagen retiniana es esencialmente bidimensional. Describir el mundo en su magnífica plenitud tridimensional requiere que el sistema visual efectúe algunas inferencias bastante complejas, inferencias que en su mayoría se hacen sin que caigamos en la cuenta de ello. Así, por ejemplo, la superficie cosenoidal de nuestra ilustración es, como la imagen retiniana de la misma, bidimensional. Y, sin embargo, aparece, con toda su fuerza, tridimensional. La apariencia de profundidad o de relieve es enteramente inferida, o, para decirlo de otro modo, es el resultado de una alucinación. Lo cual debería producirnos alguna inquietud: si, como estoy sugiriendo, tales alucinaciones no son la excepción, sino la regla, y si de hecho acompañan necesariamente a la percepción visual, ¿cómo puede justificar uno su fe en la percepción? ¿Cómo es posible, no obstante, que, en general, el ver conlleve el creer?

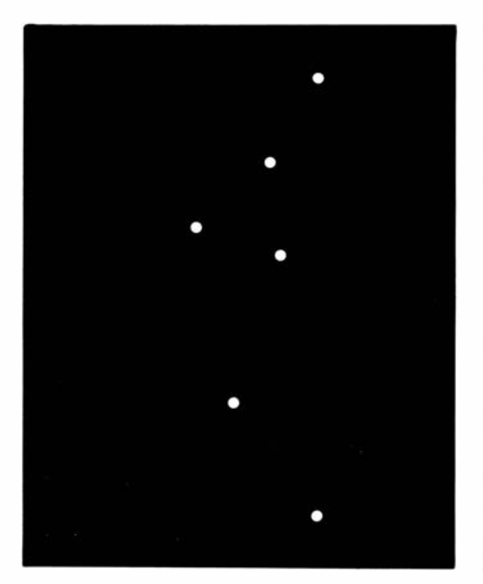
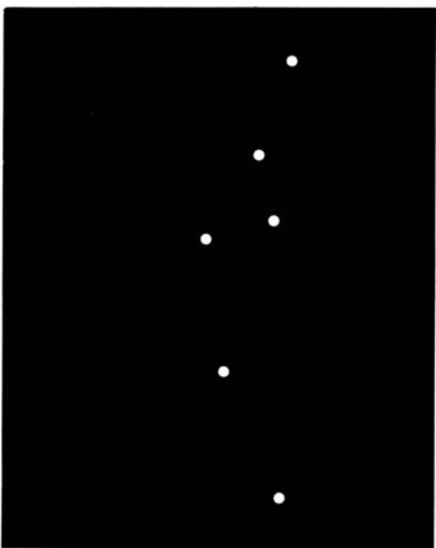
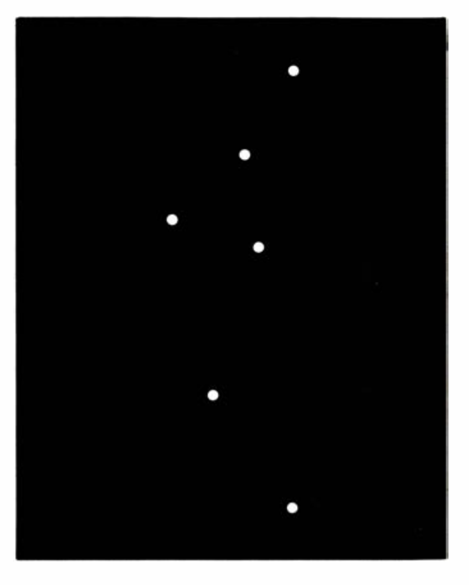
Por tanto, lo que se requiere para

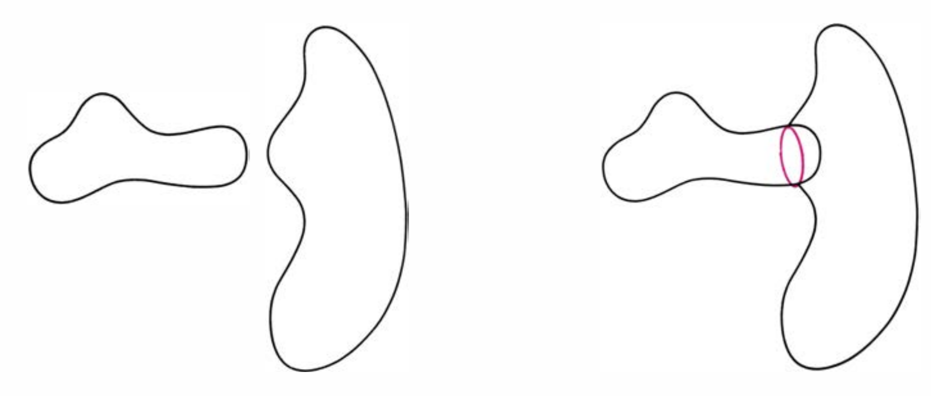
2. SE VE UNA ESFERA ROTANTE cuando se muestran en rápida sucesión los tres conjuntos de motas dispersas aquí representados. El sistema visual parece adoptar la más rígida interpretación tridimensional de las motas en movimiento que resulte compatible con las proyecciones bidimensionales.

comprender la visión es explicar por qué tales inferencias visuales suelen comportar una relación no arbitraria con el mundo real. Una línea de investigación prometedora empieza por la observación de que el mundo visible, lejos de ser enteramente caótico, obedece a ciertas leyes y exhibe numerosas regularidades. Si el sistema visual está adaptado para explotar esas leyes y regularidades en su organización e interpretación de las imágenes retinianas, y si está constreñido de algún modo a preferir la interpretación más creíble, dados la imagen y un conocimiento de esas leyes y regularidades, entonces quizá sea posible entender cómo las propias alucinaciones visuales de uno mismo comportan una relación no arbitraria y hasta útil con el mundo externo.

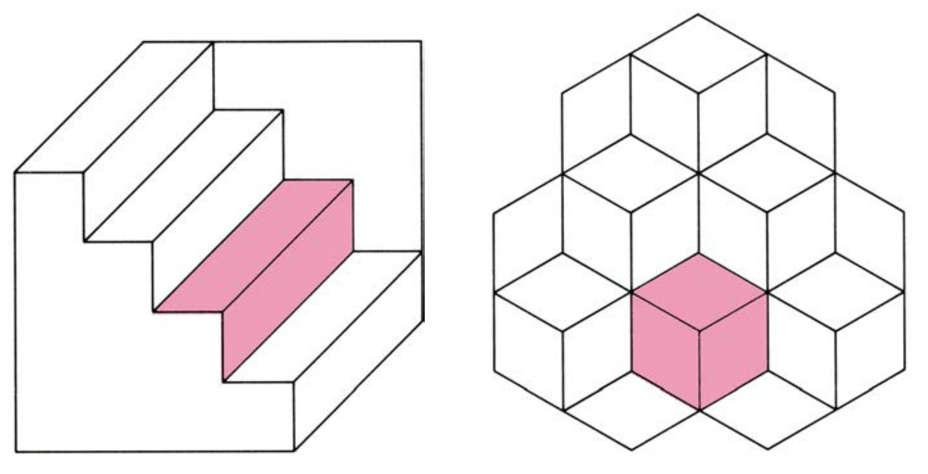
Un ejemplo especialmente claro de este enfoque lo constituye la investigación sobre la percepción visual del movimiento llevada a cabo por Shimon Ullman, del Instituto de Tecnología de Massachusetts. Ullman ha explorado la notable capacidad que posee el sistema visual humano para percibir con corrección la estructura tridimensional y el movimiento de un objeto a partir tan sólo del desplazamiento de su proyección bidimensional, capacidad que Hans Wallach y Donald N. O'Connell, del Swarthmore College, llaman efecto cinético-profundo. Con un ejemplo: si en el interior de una habitación a oscuras se hace rotar un balón playero transparente sobre cuya superficie se hayan montado unos cuantos farolillos luminosos distribuidos al azar, percíbese inmediatamente la correcta disposición esférica de las luces [*véase la figura 2*]. Al parar la rotación, cesa también la percepción de la disposición esférica. ¿Cómo es que se ve la correcta estructura tridimensional, siendo así que con la proyección retiniana bidimensional móvil son compatibles infinitas estructuras tridimensionales? Ullman demostró matemáticamente que si el sistema visual se aprovechaba de las leyes de la proyección y si sacaba partido del hecho de que el mundo contiene objetos rígidos, entonces sólo podía obtenerse, en principio, una única y correcta interpretación. En particular, mostró que para resolver el problema bastaban tres visiones de cuatro farolillos no coplanarios. La clave está en que una regla de inferencia, basada en una ley (la ley de la proyección) y una regularidad (a saber, el hecho de que el mundo contiene objetos rígidos), capacitan al sistema visual para realizar una interpretación correcta.

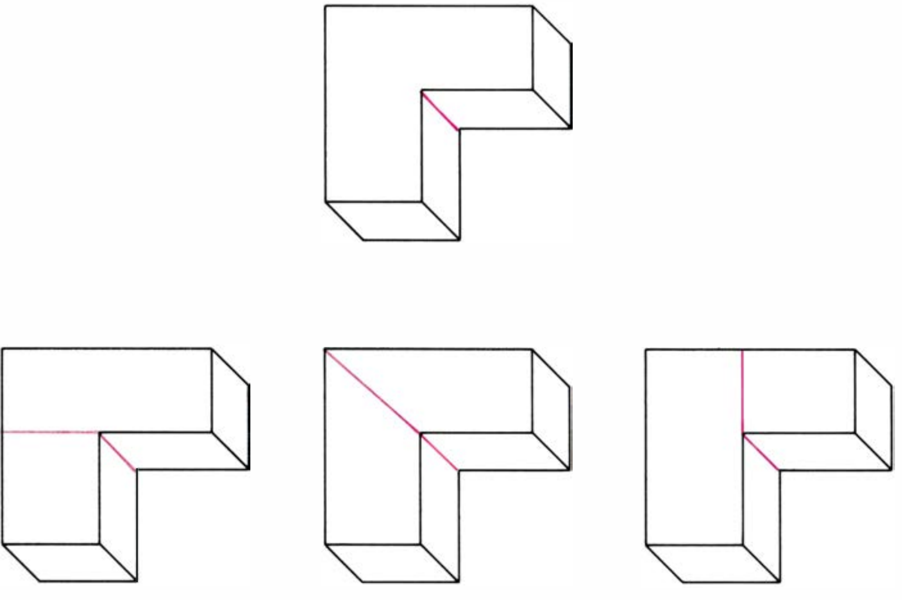
3. SE VE UNA PERSONA QUE ANDA al mostrarse en rápida sucesión estos conjuntos de motas dispersas. En este caso el sistema visual parece adoptar la interpretación tridimensional más rígida y planaria que resulte compatible con los movimientos bidimensionales de las motas. La demostración se basa en un experimento realizado por Gunnar Johansson, de la Universidad de Uppsala, en el que a una persona se le pusieron bombillitas encendidas en las principales articulaciones (hombro, codo, muñeca, cadera, rodilla y tobillo) y se filmó la escena mientras el portador se paseaba por un cuarto oscuro.



4. LA TRANSVERSALIDAD es un tipo de regularidad comúnmente observable en el mundo exterior y que subyace a una explicación unificada de varias ilusiones ópticas. Según la regla de la transversalidad (tal como la han definido Whitman A. Richards y el autor), cuando dos superficies cualesquiera se interpenetran al azar, se encuentran en una discontinuidad cóncava, indicada aquí por el contorno en color.



5. REGLA DE PARTICION, basada en la regularidad de la transversalidad. Se demuestra con ayuda de estas dos figuras, las cuales se invierten al mirarlas detenidamente. En ambos casos, los límites aparentes de las distintas partes de la forma percibida cambian al convertirse la "figura" en "fondo" y viceversa. Por ejemplo, en el caso de la ilusión de la escalera reversible (*izquierda*), publicada por H. Schröder en 1858, los dos paneles coloreados, que en una visión parecen partes de un escalón, semejan partes de dos escalones adyacentes al invertirse la escalera. Asimismo, en la ilusión del montón de cubos (*derecha*) los tres paneles romboidales coloreados pueden verse como las caras de un solo cubo o, al invertirse la figura, como las de tres cubos diferentes. Cambia, también, el número de escalones o de cubos percibidos.



6. ESTOS BLOQUES DE FORMA ACODADA muestran que la regla de la partición de las figuras en sus discontinuidades cóncavas es conservadora. La regla no da un contorno cerrado en el bloque superior porque parece posible percibir tres particiones diferentes, según lo ilustran los bloques inferiores.

Pero, llegados aquí, surge una paradoja: la misma precisión matemática que prueba que la regularidad de la rigidez es suficiente en principio para interpretar el balón playero en rotación prueba también que la regularidad de la rigidez es de suyo insuficiente para interpretar semejante exhibición. El primero en idear esa exhibición fue Gunnar Johansson, de la Universidad de Uppsala, quien la concibió como una muestra de lo que él denomina movimiento biológico. Johansson puso unas bombillitas encendidas en las principales articulaciones de una persona y filmó las luces en movimiento mientras el portador se paseaba por un cuarto oscuro. Cada fotograma parece sólo una colección de motas blancas distribuidas al azar sobre un fondo negro. En cambio, al ponerse en movimiento la cinta, se ve inmediatamente la correcta estructura tridimensional de las motas y se reconoce que hay un paseante invisible [*véase la figura 3*].

Cuando mi colega Bruce E. Flinchbaugh, que está ahora en los Laboratorios Bell, y yo abordamos este problema, nos intrigaba que pudiera verse la correcta estructura tridimensional aunque, a tenor de los resultados de Ullman, faltara la información apropiada para ello. Para inferir una correcta estructura tridimensional partiendo de la base de la regularidad de la rigidez es necesario tener tres instantáneas de, por lo menos, cuatro puntos no coplanarios que estén en una configuración rígida. Por otra parte, en las exhibiciones de movimiento biológico, sólo están conectados rígidamente, a lo sumo, pares de puntos, tales como el tobillo y la rodilla o la rodilla y la cadera. Lo que precisamente no existen son cuaternas rígidas de puntos.

La regularidad de la rigidez es, por tanto, insuficiente de por sí, llevándonos a preguntar: ¿qué ulterior regularidad explota el sistema visual? Tras varios planteamientos fallidos, se nos ocurrió que la solución tal vez fuese una regularidad anatómica. En la mayoría de los animales los miembros que soportan el peso están obligados, por la construcción de sus junturas, a girar en un plano único cuando se trata de caminar al paso normal. A eso lo llamamos la regularidad de la planaridad.

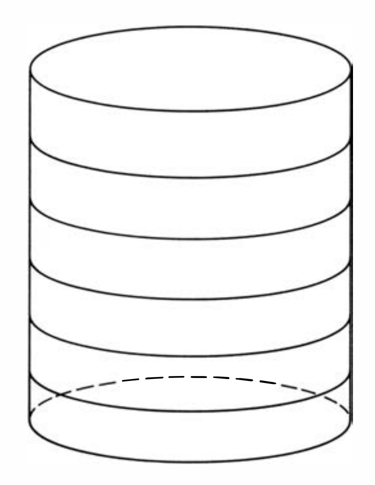
De hecho, la regularidad de la planaridad es suficiente para interpretar con corrección exhibiciones del movimiento biológico de caminar al paso. La correcta estructura tridimensional puede inferirse bien de tres instantáneas de dos puntos que se muevan con rigidez

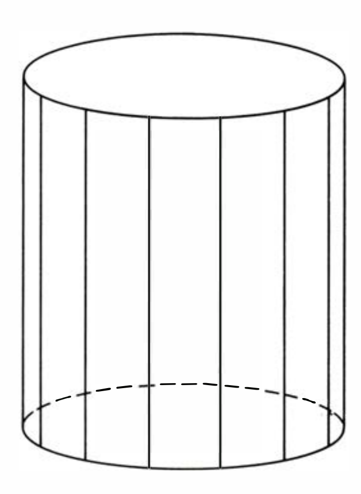
en un plano, o bien de dos instantáneas de tres puntos (como tobillo, rodilla y cadera) que formen pares rígidos y se muevan en un mismo plano. Estos resultados concuerdan perfectamente con la observación de Johansson de que sólo precisa verse dos o tres fotos de sus filmes para que los sujetos perciban correctamente el movimiento biológico. Un resultado adicional es que no sólo pueden interpretarse correctamente todos los movimientos tridimensionales regidos por la regularidad de la planaridad, sino que siempre que se halle una interpretación para el movimiento de imagen basado en la regularidad de la planaridad o en la regularidad de la rigidez la interpretación será también correcta.

En resumen, la probabilidad de que la interpretación sea errónea es nula si se supone una resolución infinita en la imagen, o una probabilidad muy pequeña si la resolución no es perfecta del todo. De ahí que las estructuras no rígidas no puedan hacerse pasar por rígidas ni los movimientos no planarios puedan ser mal construidos como planarios. Una vez más, las leyes y las regularidades prueban estar en el centro de la explicación de cómo el sistema visual logra una única y correcta interpretación de una imagen retiniana.

Volvamos ahora a la superficie cosenoidal. Su principal interés reside en que nos muestra el sistema visual organizando figuras por partes, organización que es sumamente útil cuando se trata de reconocer un objeto a partir de su forma. La superficie cosenoidal revela también que el invertir una figura puede alterar su organización. ¿Es, pues, el sistema visual caprichoso en su organización? No parece probable. Pero si no lo gobierna el capricho, han de hacerlo reglas a las que se ajuste al definir las partes. Y si las reglas no pueden ser arbitrarias, tendrán que basarse en alguna ley o regularidad que se dé en el mundo externo.

Esta línea de razonamiento nos llevó a Whitman A. Richards, del MIT, y a mí a buscar una ley o regularidad que motivara una serie de reglas del reparto de superficies. La regularidad que nos pareció fundamental atañe a la transversalidad: si se hace que dos superficies de arbitraria configuración se interpenetren recíprocamente al azar, siempre se juntan en un contorno de discontinuidad cóncava de sus planos tangentes [*véase la figura 4*]. Aunque la regularidad de la transversalidad pueda sonar a esotérica, se trata de algo muy corriente en la experiencia cotidiana. En una bebida sin alcohol, por ejemplo, una paja forma, allí donde se junta con la superficie del líquido, una discontinuidad cóncava y circular. La velita en la tarta de cumpleaños, las puntas del tenedor en un trozo de filete, el cigarrillo entre los labios... son otros tantos ejemplos de esta ubicua regularidad.

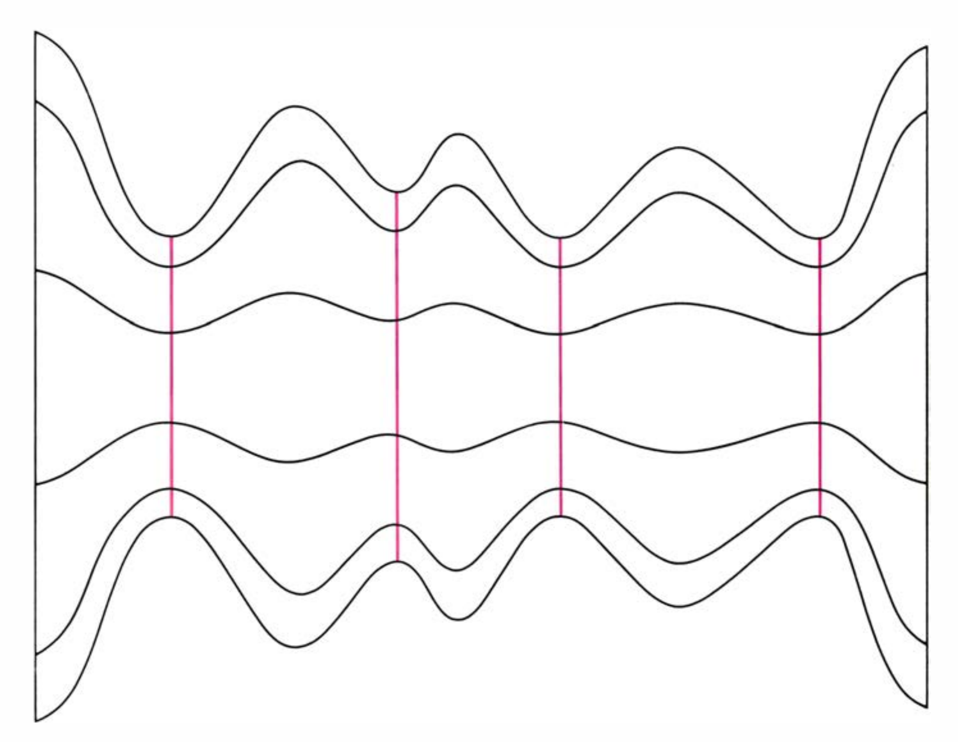




7. LAS LINEAS DE CURVATURA son fáciles de representar sobre un vaso cilíndrico ideal. Las líneas de curvatura máxima (*izquierda*) son círculos; las líneas de curvatura mínima (*derecha*) son rectas.

Sobre la base de la regularidad de la transversalidad puede proponerse una primera regla para el reparto de una superficie: divídese una superficie en partes siguiendo todos los contornos de discontinuidad cóncava. Esta regla no sirve en el caso de la superficie cosenoidal, porque ésta es enteramente lisa. La regla debe primero generalizarse un poco, como lo haremos más abajo. Pero en su forma restringida puede dilucidar varias demostraciones perceptuales bien conocidas.

Por ejemplo, la regla hace la predicción obvia de que las partes de la escalera mostrada en la figura 5 son sus escalones, con un escalón entre dos sucesivas líneas de discontinuidad cóncava de la escalera. La regla predice también algo no tan obvio: si la escalera experimenta una inversión perceptual, de modo que lo "saliente" se convierta en "fondo" y viceversa, entonces los límites de los escalones tienen que cambiar. Esta conclusión se sigue porque sólo las discontinuidades cóncavas definen los límites de los escalones, y lo que parece una concavidad desde un lado de una superficie parece una convexidad desde



8. DELIMITACIONES DE PARTES según las define la regla generalizada de partición de una superficie lisa. Se han representado por trazos de color sobre esta superficie de arbitraria configuración. Las líneas negras son las de máxima curvatura, cuyos mínimos dan origen a los trazos de partición coloreados.

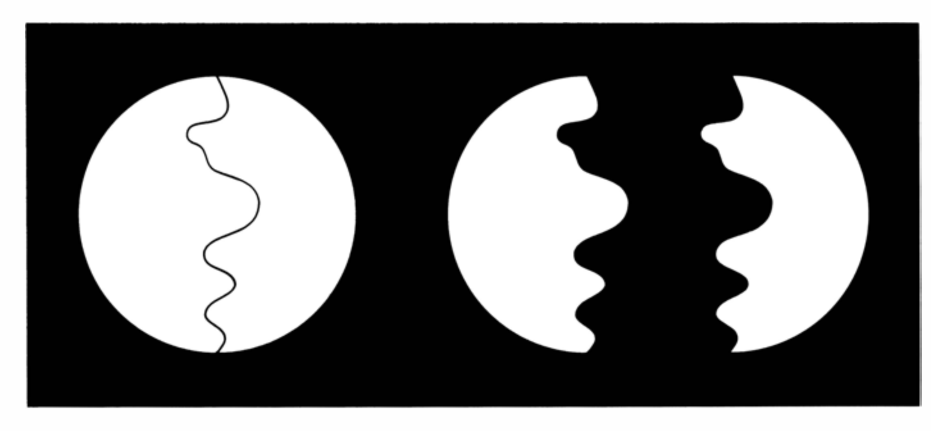
el otro. Así, al invertirse la escalera, las discontinuidades convexas y las cóncavas han de invertir sus papeles, fijando nuevos límites a los escalones. Esta predicción puede comprobarse mirando el escalón coloreado en sus dos caras: cuando la escalera aparece vuelta del revés se advierte cómo los paneles coloreados no están ya en un solo escalón, sino en dos escalones adyacentes.

La predicción se puede confirmar con una demostración más complicada, como lo es la prueba de los cubos amontonados que vemos en la misma ilustración. Las tres caras coloreadas que al principio parecen estar en un cubo aparecen en tres cubos cuando se invierte la figura.

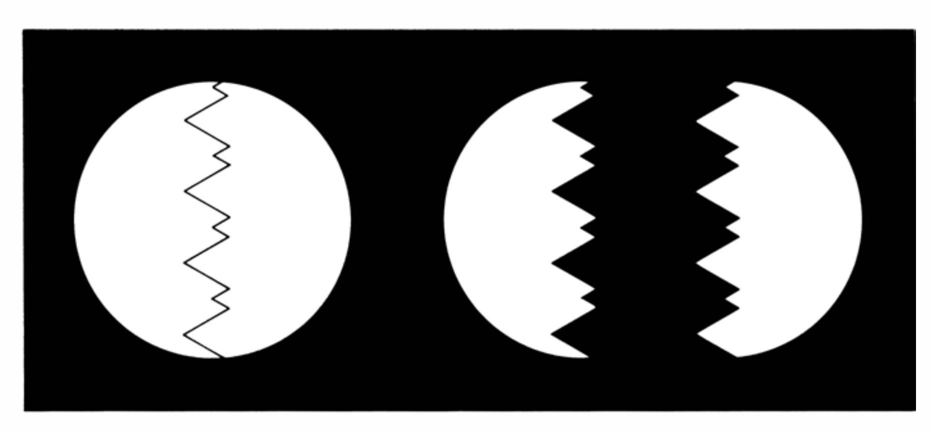
De esta simple regla de reparto síguese una ulterior predicción. Si la regla no define una única partición de alguna superficie, entonces el modo apropiado de dividir la superficie en partes ha de ser perceptualmente ambiguo (a menos que haya reglas adicionales que eliminen la ambigüedad). Una confirmación clara de esta predicción puede verse con respecto al bloque de forma acodada representado en la figura 6. La única discontinuidad cóncava es la línea vertical en el pliegue del codo. Por consiguiente, la regla no define una única partición del bloque. Perceptualmente hay tres modos plausibles de dividir el bloque en partes. Los tres cuentan con el contorno definido por la regla de partición, pero lo completan siguiendo diferentes caminos.

Esta simple regla de partición lleva a interesantes aclaraciones de la percepción de la forma. Sin embargo, para explorar la superficie cosenoide y otras superficies lisas, la regla debe generalizarse. Lo cual requiere una breve digresión sobre la geometría diferencial de las superficies, a fin de que se entiendan tres importantes conceptos: normal a la superficie, curvatura principal y línea de curvatura. Por suerte, aunque estos conceptos son muy técnicos, se les puede dar una fácil caracterización intuitiva.

La normal a la superficie en un punto dado de una superficie puede representarse con una aguja de una unidad de longitud clavada perpendicularmente a esa superficie en dicho punto. A todas las normales a la superficie en todos los puntos de una superficie –algo así como las espinas de un erizo de mar– se les llama colectivamente "campo de normales a la superficie". De ordinario, sobre una superficie hay dos posibles campos de normales a la superficie: según apunten las normales hacia fuera o hacia dentro. Por ejemplo, una esfera puede tener las normales a la superficie apuntando todas radialmente hacia fuera, como espinas, o todas hacia dentro, hacia su centro. Admitamos, por convenio, que el campo de normales a la superficie apunta siempre al interior de la figura. Así, una pelota de béisbol tendrá las normales hacia dentro, mientras que una burbuja bajo el agua tendrá las normales hacia fuera. El invertir la opción de figura y fondo sobre una superficie implica el concomitante cambio de las normales a la superficie. Una inversión del campo de las normales a la superficie trae consigo un cambio de signo de cada una de las curvaturas principales en cada punto de la superficie.



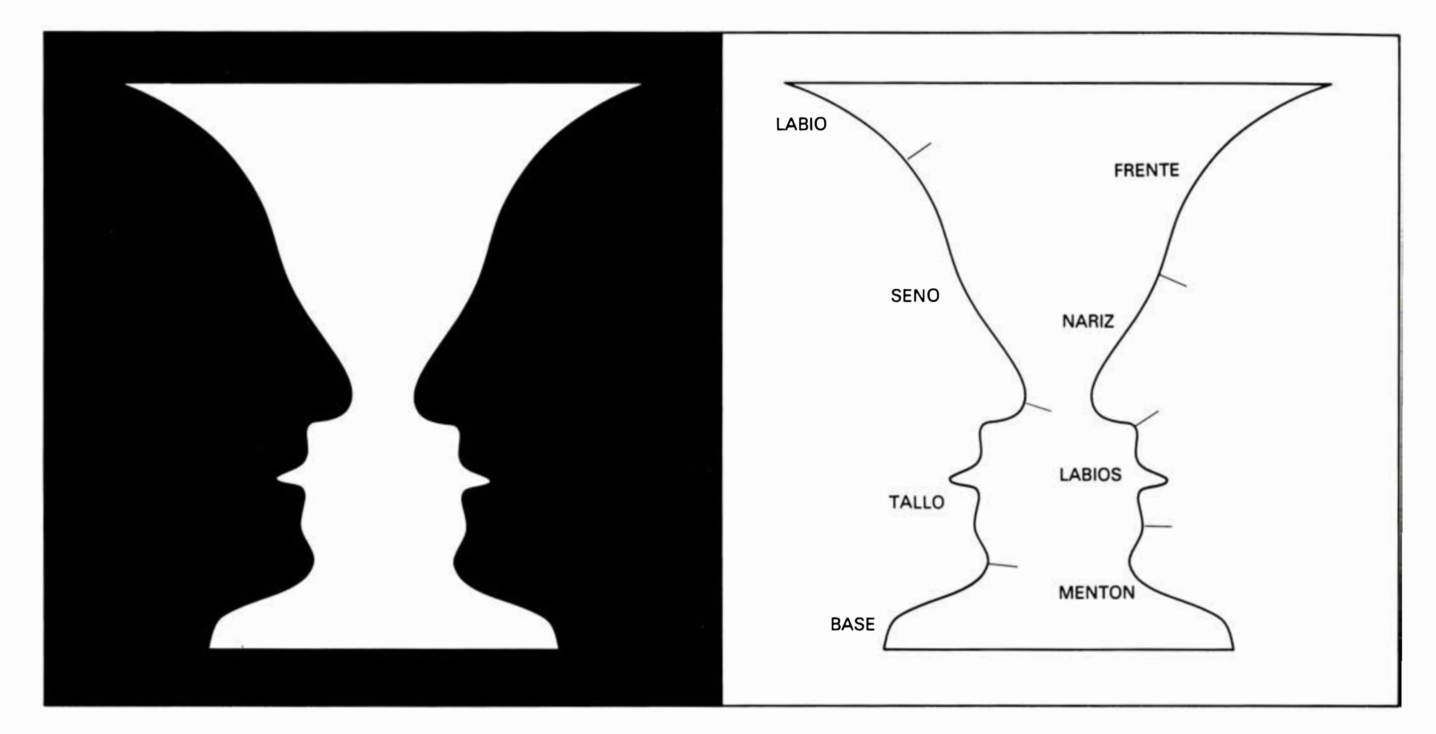
9. CURVA PLANA REVERSIBLE, construida por Fred Attneave, de la Universidad de Oregón, garabateando un trazo aleatorio a través de un círculo y separando las dos mitades. Véase cómo la forma aparente del contorno resultante depende de qué lado del trazo se perciba como "figura" o como "fondo".



10. UNA FIGURA REVERSIBLE SIMILAR puede obtenerse con una curva plana de curso no suave. El contorno dentellado resultante parecerá una cadena de montañas altas y bajas alternando, o, si se decide invertir la asignación de figura y fondo, una cadena de montañas altas y con dos picos cada una.

A menudo importa conocer no sólo la normal a la superficie en un punto, sino también cómo se curva la superficie en ese punto. El matemático suizo del siglo XVIII Leonhard Euler descubrió que en todo punto de cualquier superficie hay siempre una dirección en la que la superficie se incurva el mínimo, y una segunda dirección, siempre perpendicular a la primera, en la que la superficie se incurva el máximo. (Tratándose de un plano o de una esfera, la curvatura de la superficie es idéntica en todas direcciones y en cualquier punto.) A esas dos direcciones se las ha llamado direcciones principales, y a las correspondientes curvaturas de la superficie, curvaturas principales. Partiendo de algún punto y moviéndose siempre en la dirección de la mayor curvatura principal, se traza una línea de curvatura máxima. Moviéndose, en cambio, en la dirección de la menor curvatura principal, se traza una curva de curvatura mínima. En la superficie de un vaso de mesa, la familia de las líneas de curvatura máxima es una serie de círculos que ciñen el vaso. Las líneas de curvatura mínima son líneas rectas que recorren la longitud o altura del vaso [*véase la figura 7*].

Con estos conceptos en las mientes, es fácil hacer extensiva la regularidad de la transversalidad a las superficies lisas. Supóngase que en cualquier sitio donde una superficie ofrezca una discontinuidad cóncava la aminora uno un poco disponiendo sobre ella una piel tensa. Entonces, una discontinuidad



11. ILUSION DE LA COPA Y LOS ROSTROS, ideada por Edgar Rubin hacia 1915. Puede verse una copa o dos perfiles humanos (*izquierda*). Si se considera que la figura son los rostros, repartiéndola por referencia a los mínimos de curvatura se dividirán los perfiles en tramos correspondientes a frentes, narices, labios y mentones; si se toma como figura la copa, definiendo los límites de las partes por los mínimos de curvatura se dividirá el contorno en borde o labio, seno, tallo y base (*a la derecha*). En ambos casos las líneas normales principales se han dispuesto hacia el interior de la figura.

cóncava se convierte, poco más o menos, en un contorno en el que la superficie tiene localmente la máxima curvatura negativa. Precisando más, la versión generalizada de la transversalidad sugiere la siguiente regla general de partición para las superficies: una superficie se divide en partes en los mínimos negativos de cada curvatura principal según la familia de líneas de curvatura que tiene asociada [*véase la figura 8*].

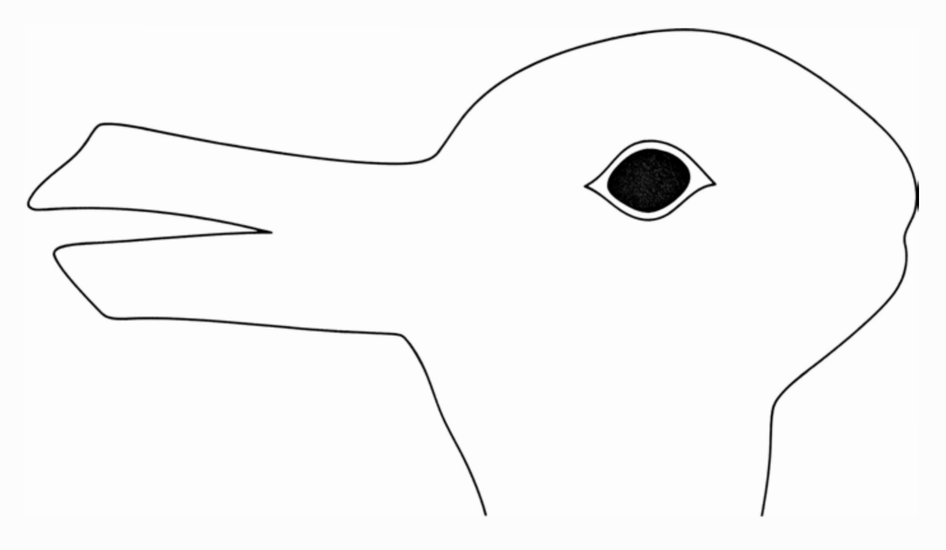
Esta regla divide la superficie cosenoidal según los contornos circulares rojos. Explica también por qué las partes cambian al girarse la página de arriba abajo: el sistema visual invierte entonces su atribución de figura y fondo sobre la superficie (quizá debido a una preferencia por una interpretación que pone el objeto por debajo del punto de vista del observador en lugar de por encima de él). Cuando figura y fondo se invierten, sucede lo mismo con el campo de las normales a la superficie, de acuerdo con el convenio antes mencionado. Cálculos simples demuestran, empero, que cuando las normales se invierten cambian también de signo las curvaturas principales. A resultas de lo cual, los mínimos de las curvaturas principales han de convertirse en máximos y viceversa. Y como los mínimos de las curvaturas principales se utilizan para delimitar las partes, síguese que también estos límites de las partes han de moverse. En suma, las partes parecen cambiar porque la regla de partición, motivada por la regularidad de la transversalidad, usa los mínimos de las curvaturas principales, y porque estos mínimos cambian de sitio en la superficie al invertirse figura y fondo.

Para decirlo en pocas palabras, la regularidad de la transversalidad proporciona una unidad subyacente a los relatos explicativos de la percepción de partes en superficies lisas y en superficies accidentadas. Y subyace asimismo a una explicación de otra clase muy conocida de ilusiones ópticas: las curvas planas reversibles. Un buen ejemplo de este fenómeno es la figura reversible ideada por Fred Attneave, de la Universidad de Oregón [*véase la figura 9*]. Halló éste que con sólo garabatear una línea a través de un círculo y separar los dos semicírculos así creados se pueden obtener dos perfiles muy diferentes y contrapuestos. Está claro, como lo hace notar Attneave, que el aspecto del contorno depende de qué lado se considere que es parte de la figura y no de alguna familiaridad previa con ese contorno.

¿Cómo explica este fenómeno la regularidad de la transversalidad? La respuesta consta de tres pasos: (1) una proyección de la regularidad de la transversalidad de las tres a las dos dimensiones, (2) una breve digresión sobre la geometría diferencial de las curvas planas y (3) la formulación de una regla de reparto de las curvas planas.

La versión bidimensional de la regularidad de la transversalidad es similar a la versión tridimensional. Si se hace que dos superficies arbitrariamente configuradas se interpenetren al azar, en cualquier proyección bidimensional de la superficie compuesta resultante siempre se encontrarán la una con la otra en cúspides cóncavas. Parafraseando libremente diremos que en toda silueta se forman cúspides cóncavas en los puntos donde termina una parte y empieza otra. Lo cual sugiere la siguiente regla de partición de las curvas planas: una curva plana se divide en partes en las cúspides cóncavas. Esta regla no puede aplicarse a la demostración de Attneave porque tal demostración se basa en un contorno que es todo él suave. La regla debe generalizarse una vez más. Sin embargo, en su forma no generalizada, puede dar razón de una versión de la prueba de Attneave que no presente sólo perfiles suaves.

En la figura 10, el mismo contorno mellado puede parecer, bien una cadena de montañas alternativamente altas y bajas, o bien, para la asignación figura-fondo inversa, una cadena de altas montañas con doble pico o cima cada una. El perfil se divide en partes de forma distinta al invertirse la figura y el fondo porque la regla de división, para delimitar las partes, usa sólo las cúspides cóncavas. Lo que es una cúspi-



12. ILUSION DEL ANIMAL REVERSIBLE, que no implica la inversión de figura y fondo. Conforme a ello, los límites de partes definidos por los mínimos de curvatura no cambian de posición al alterarse la interpretación. Las orejas del conejo pasan a ser el pico del pato sin que se produzca movimiento.

de cóncava cuando un lado del perfil es figura ha de pasar a ser cúspide convexa cuando es figura el otro lado, y viceversa. Se da un paralelismo entre este ejemplo y el de la escalera reversible de que nos ocupamos más arriba.

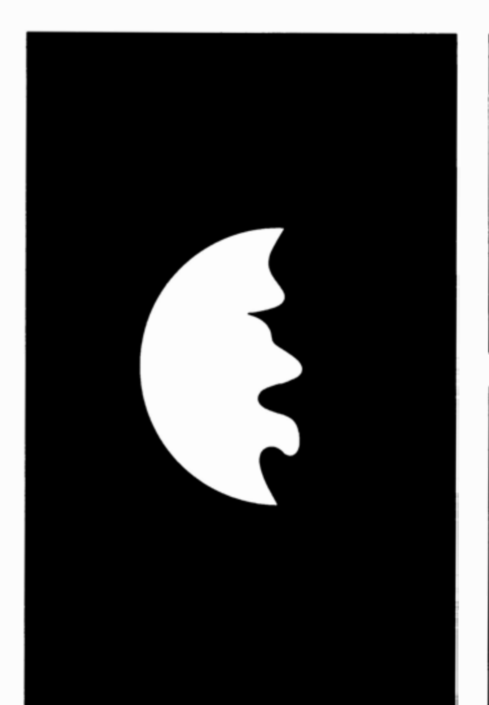
Antes de generalizar la regla para los contornos suaves, repasemos brevemente dos conceptos de la geometría diferencial de las curvas planas: la normal y la curvatura principales. La normal principal a un punto de una curva se puede concebir como una aguja de una unidad de longitud clavada verticalmente en ese punto de la curva. El conjunto de todas las normales principales a todos los puntos de una curva forma un campo de normales principales. Por lo común, hay dos posibles campos de normales principales, uno por cada lado de la curva. Convengamos en suponer que el campo de normales principales se elige siempre para apuntar hacia el "lado figura" de la curva. Invertir respecto a una curva la elección de figura y fondo implica un cambio concomitante del campo de normales principales. Importa que advirtamos que, por el convenio de forzar a las normales principales a que apunten hacia la figura, las partes cóncavas de una curva tienen curvatura negativa y las partes convexas curvatura positiva.

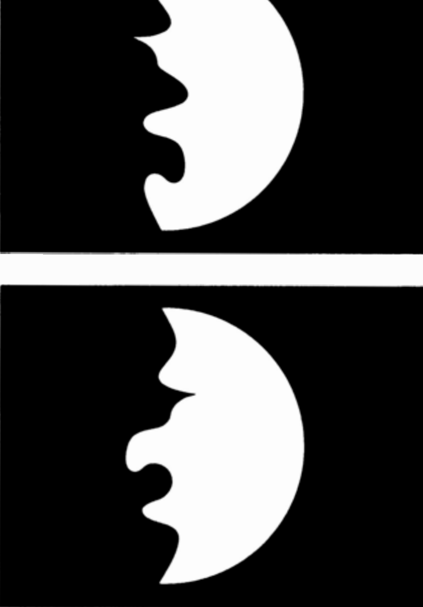
Ahora es ya fácil tarea generalizar la regla de división para las curvas planas. Supongamos que dondequiera que una curva tenga una cúspide cóncava suavizamos la curvatura un poco. Entonces, una cúspide cóncava se convierte en un punto de curvatura negativa que tiene, localmente, el mayor valor absoluto de curvatura. Esta observación lleva a la siguiente regla general de división: una curva plana se divide en partes en los mínimos de curvatura negativa.

Podemos entonces explicar por qué las dos mitades del círculo de Attneave parecen tan diferentes. Al invertirse figura y fondo, el campo de las normales principales se invierte también de acuerdo con lo convenido, y al invertirse las normales principales ha de cambiar de signo la curvatura en todos y cada uno de los puntos de la curva. Concretamente, los mínimos de curvatura han de convertirse en máximos y viceversa. Este reposicionamiento de los mínimos de curvatura lleva a una nueva división de la curva según la regla de división. En una palabra, la curva parece diferente por estar organizada en tramos fundamentalmente diferentes, en unidades distintas. Nótese que si se elige el definir las delimitaciones de las partes atendiendo a inflexiones, o por máximos y mínimos de curvatura a la vez, entonces no suelen cambiar los tramos cuando se invierten figura y fondo.

Un ejemplo claro de dos reparticiones muy diferentes de una curva puede verse en la famosa ilusión de los rostros y la copa propuesta hacia 1915 por Edgar Rubin [*véase la figura 11*]. Si se considera que la figura son los rostros, los mínimos de curvatura dividen la curva en tramos que corresponden a frentes, narices, bocas y mentones. En cambio, si se toma por figura la copa, se reposicionan los mínimos, que dividen la curva en tramos distintos, correspondientes a los perfiles de la base, el tallo, el seno y el borde o labio de la copa. Probablemente no es pura casualidad que a las partes definidas por los mínimos se les asignen con frecuencia etiquetas verbales.

Se han ideado exhibiciones que, como en el caso de la ilusión rostros-copa, permiten más de una interpretación de un único contorno pero que no implican una inversión de figura y fondo. Ejemplo popular de ello es la





13. PRUEBA EL TEST DE LA MEDIALUNA que los juicios sobre parecidos de formas pueden predecirse correctamente mediante la regla de partición según los mínimos de curvatura. A una primera mirada, la medialuna inferior de la derecha se ve más parecida a la única medialuna de la izquierda que la superior de la derecha. Sin embargo, mirando con mayor atención, se advierte que el perfil de la medialuna superior de la derecha es idéntico al de la medialuna de la izquierda, mientras que el perfil de la medialuna inferior de la derecha se ha invertido especularmente y tiene dos partes intercambiadas.

ilusión conejo-pato [*véase la figura 12*]. Como tales ilusiones no implican una inversión figura-fondo y, por tanto, nunca cambian en ellas de posición los mínimos de curvatura, la regla de repartición debe predecir que los límites de las partes serán idénticos para las interpretaciones de cada uno de los contornos. Esta predicción es fácil de confirmar: lo que es, por así decirlo, una oreja del conejo pasa a ser una parte del pico en el pato.

Si el sistema visual humano obedece la regla de los mínimos para el reparto de curvas, podría suponerse que predijera algunos juicios de similitud de formas. Un caso en el que esta predicción es contraria a lo intuitivo se recoge en la figura 13. Mírese un momento la medialuna que hay en el recuadro izquierdo de la ilustración. Mírense luego rápidamente las dos mediaslunas de la derecha y decídase cuál de ellas se parece más a la primera. En un experimento efectuado con varias figuras similares, Aaron F. Bobick, del Instituto de Tecnología de Massachusetts, y yo comprobamos que casi todos los sujetos elegían como la más parecida la medialuna del recuadro inferior derecha. Y, sin embargo, si se observa otra vez se hallará que el perfil de la medialuna superior derecha es idéntico al de la medialuna de la izquierda, sólo que con su relación figura-fondo invertida. El perfil delimitante de la medialuna inferior se ha invertido especularmente y se han cambiado dos partes definidas por mínimos de curvatura. ¿Por qué, de hecho, la medialuna inferior sigue pareciendo más semejante? La regla de los mínimos da una respuesta sencilla. El contorno inferior, que no está invertido en cuanto a figura-fondo respecto al contorno primero, tiene las mismas delimitaciones de partes. El contorno superior, que está invertido en cuanto a figura-fondo respecto al primero, tiene, en cambio, diferentes delimitaciones de partes.

Resumiendo: la visión es un proceso activo cuya función consiste en inferir descripciones útiles del mundo a partir de los cambiantes haces luminosos que inciden en las retinas. Estas descripciones sólo son fiables en tanto en cuanto los procesos inferenciales que las construyen exploten regularidades que se dan en el mundo visual, tales como la rigidez, la planaridad y la transversalidad. El descubrimiento de regularidades importantes y la investigación matemática de su poder guiador de las inferencias visuales son direcciones prometedoras para el investigador que trate de comprender la visión humana.

Juegos de ordenador

Las máquinas estado-finitas, modelos minimales de ratoneras, de ribosomas y del alma humana

Brian Hayes

Los más potentes ordenadores carecen por igual de soporte físico o "cacharrería" (eso que llaman *hardware*) y de soporte lógico o programas (*software*); se construyen con elementos puramente conceptuales. La más célebre de estas máquinas abstractas es la ideada en 1936 por el matemático inglés Alan Mathison Turing, y es capaz de más que pueda cualquier ordenador construido con microcircuitos de silicio; la de Turing computa todo lo computable. Hay una categoría de computadoras conceptuales próxima a la de Turing que, si bien carecen de la universalidad de ésta, no son por ello menos interesantes. Son las llamadas máquinas de número finito de estados, y establecen las especificaciones mínimas a cumplir por un ordenador digital que pueda funcionar.

El rigorismo matemático preciso para definir adecuadamente la máquina con finitud de estados estaría aquí fuera de lugar. Podemos, empero, aclarar la naturaleza de tal noción mediante unos cuantos ejemplos. Al ir yo buscando máquinas de tipo finito, encontré un excelente espécimen en la estación de Metro de Lexington Avenue, Nueva York. Se trataba de un molinete de acceso y cobro, de los antiguos, que, en vez de el clásico trípode de acero macizo de hoy, disponía de cuatro travesaños de roble, que un río de manos y caderas han desgastado y pulido.

Dos son los estados del molinete: bloqueado y franco. Supongámoslo en estado de bloqueo, con lo que no podremos hacer girar sus aspas. La inserción de una ficha de pago por la rendija de cobro modifica el estado del mecanismo interno en modo tal que sus brazos quedan ahora libres para girar. Dicho de otra forma, la ficha induce una transición al estado franco. Vemos esquematizadas las transiciones en la figura 1. Los estados del sistema están representados por nodos (recuadros) y, las transiciones, por arcos (flechas) que los conectan.

En un análisis de tipo finito del molinete, las dos posibles entradas admitidas por el sistema serían la inserción de una ficha y el empuje sobre uno de sus brazos. La respuesta de la máquina depende tanto de la señal de entrada que se le aplique como del estado en que se encuentre en ese momento. Mientras el molinete no sea alimentado con una ficha no podremos dar una vuelta en metro. Insertar una ficha estando el paso ya franco es igualmente fútil, si bien en un sentido algo diferente. La segunda ficha es aceptada, pero no influye sobre el estado de la máquina. Es admitida sólo una persona; seguidamente el molinete queda bloqueado. Se aceptan igualmente tres o cuatro fichas, pero sólo se da viaje para uno. Los escépticos tienen derecho a exigir más pruebas antes de aceptar que todas las fichas que puedan insertarse a continuación de la primera carecerán de efecto, pero la provisión de fichas para nuevos experimentos queda a su cargo.

La razón por la que el molinete no puede dar más de un viaje al recibir más de una ficha es que carece de elementos para contar las recibidas. Su única forma de memoria es muy rudimentaria: al pasar de uno a otro estado "se acuerda" de si la más reciente señal de entrada recibida fue una ficha o un empujón sobre uno de sus brazos. Todas las entradas precedentes quedan "olvidadas". Vale la pena observar que su olvidadizo carácter nunca va en perjuicio de la compañía del metropolitano. La cosa podría ser peor: podría diseñarse un molinete que cambiase de estado cada vez que recibiera una ficha, independientemente de su estado actual, en cuyo caso al insertar dos fichas seguidas no podríamos pasar.

El molinete ilustra casi todas las propiedades fundamentales de la máquina de número finito de estados (que llamaremos, para abreviar, máquina estado-finita). Evidentemente, la máquina debe tener ciertos estados, y solamente debe existir un número finito de ellos. Puede haber entradas y salidas asociadas con cualesquiera de los estados. Tales estados han de ser "discretos", es decir, individuales y perfectamente distinguibles; las transiciones entre ellos han de ser instantáneas. Mucho depende en estas cuestiones del punto de vista: día y noche son estados discretos admitiendo que amanecer y atardecer sean procesos "instantáneos". El conjunto de estados, las entradas y las salidas constituyen la máquina completa. No se permiten otros dispositivos auxiliares; en particular no está permitido el almacenacimiento de información.

Las reglas constructivas de las máquinas con número finito de estados admiten cierto margen de variación. Hay máquinas determinísticas y máquinas no determinísticas. Hay máquinas de Moore y máquinas de Mealy. En una máquina determinística, cada entrada dada de cada estado dado produce invariablemente un mismo efecto; en una máquina no determinística pueden ser varias las posibles transiciones. En la máquina de Moore (en honor de Edward F. Moore), cada estado tiene tan sólo una salida. En la máquina de Mealy (en recuerdo de G. H. Mealy), las salidas están asociadas con las transiciones, no con los estados. Resulta, empero, que tal diversidad de arquitecturas tiene mucho de ilusoria. Cualquier tarea ejecutable por uno de los tipos de máquina estado-finita puede ser también realizada por los otros tipos, si bien el número de estados puede ser distinto. Voy a comentar aquí, sobre todo, las máquinas determinísticas de Moore, que son las de estructura más sencilla.

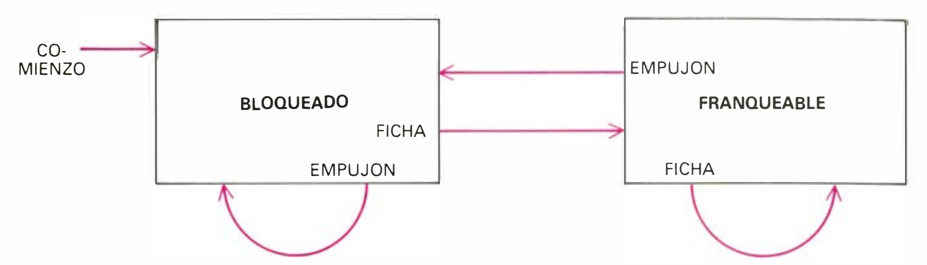
En cuanto se pone uno a buscar máquinas estado-finitas, se las encuentra por todas partes. Los ejemplos clásicos, de manual, son las máquinas de funcionamiento por monedas. Algunas máquinas de expendición automática son menos rapaces que el molinete del metro: en cuanto reciben la cantidad de dinero exacta, entran en un estado en el que todas cuantas monedas extra se le echen serán devueltas. De los dispositivos de monedas, el de máximo número de posibles estados es seguramente la tragaperras de las salas de juego de Las Vegas. Se trata, en principio, de una máquina determinística, pero la determinación de las señales de entrada (una moneda y un tirón de la palanca) capaces de provocar una transición a cierto estado final constituye, no obstante, todo un reto.

Muchos aparatos domésticos pueden considerarse máquinas estado-finitas, pero apenas si ofrecen interés. Una lavadora va pasando por una sucesión de estados (toma de agua, agitación, acla-

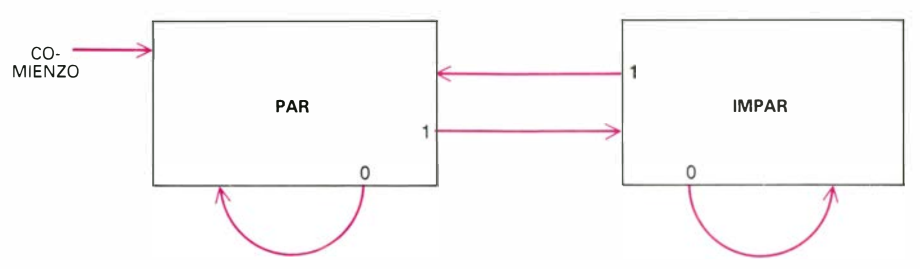
rado, centrifugado) y las escasas entradas significativas (desenchufarla, por ejemplo) tienen por lo común iguales efectos, independientemente del estado. Análogamente, un semáforo tiene un pequeño repertorio de estados, que se repiten periódica e indefinidamente. A mi juicio, la más aburrida de todas las máquinas de número finito de estados es el reloj digital. Uno de los que dan mes, fecha, y el paso de horas, minutos y segundos, tiene unos 31 millones de estados; en un año cada estado es ocupado exactamente una vez.

También una ratonera es una máquina de número finito de estados; el ratón (por lo común, para perdición suya) desencadena una transición desde el estado amartillado hasta el estado disparado. Una cerradura de combinación es una máquina de tipo finito con un gran número de posibles entradas, de las que solamente una es causa de una transición de estado. Un teléfono admite varios estados, que pueden describirse como colgado, descolgado, en espera, con tono para marcar, marcando, sonando, comunicando y estropeado. Los automóviles permiten poner vívidamente de manifiesto que el efecto de una entrada depende del estado actual del sistema. ¿Qué sucede al pisar a fondo el pedal del acelerador? Depende. ¿Está en marcha el motor? ¿Está embragado? ¿Está calzada una marcha? ¿Está echado el freno de mano? La marcha puesta, ¿es hacia adelante, o hacia atrás? ¿Está abierta la puerta del garaje?

En las células vivas, el sistema molecular formado por el ribosoma y las diversas especies de ARN de transferencia opera a modo de máquina estado-finita. Las entradas son las cuatro bases nucleotídicas del ARN mensajero, designadas por las iniciales *U, A, C* y *G*. Las salidas son los 20 aminoácidos componentes de las proteínas. Para que una cadena de nucleótidos sea reconocida como señal de entrada válida para la máquina, es preciso que comience con la señal de "puesta en marcha" *AUG*. A partir de ahí, la máquina va leyendo el chorro de entrada continuamente, cambiando de estado cada vez que reconoce un codón, o terna específica de nucleótidos. Los tres codones especiales *UAA, UAG* y *UGA* son señales de detención: tan pronto se detecta uno de ellos, la máquina se detiene. Hay otros muchos sistemas biológicos donde resulta útil la representación en máquina de tipo finito. Entre los ejemplos que nos vienen a las mientes están la molécula de hemoglobina y las proteínas represoras y promotoras de las bacterias.



1. Diagrama de transiciones entre estados de un molinete de la entrada del Metro



2. Máquina de verificación de paridad

En la teología de Sto. Tomás de Aquino, el alma resulta ser una máquina de número finito de estados, maravillosamente ideada y elaborada y plenamente determinística. Se crea en estado de riesgo, a consecuencia del pecado original. Adquiere con el bautismo el estado de gracia; pero ciertos actos (la idolatría, la blasfemia, el adulterio, etcétera) inducen en ella la transición al estado de pecado. Son precisas la confesión, el arrepentimiento y la penitencia para restaurar en el alma el estado de gracia. El efecto de una señal de entrada final e irreversible, la muerte, depende críticamente del estado del alma en el instante de la muerte: en un estado de gracia, la muerte conduce a la salvación, pero en estado de pecado lleva a la condenación eterna. La máquina del alma es en realidad más compleja de lo que esta descripción sucinta sugiere. Una descripción completa tendría que distinguir entre los distintos grados de pecado (venial y mortal, actual y habitual), y tendría que dar cabida a otros posibles estados del alma (como los asociados con el limbo y el purgatorio), y admitir otras posibles entradas (el Juicio Final, por ejemplo).

En la mecánica cuántica, hasta el átomo se convierte en una máquina de tipo finito y, por consiguiente, otro tanto acontece con todo cuanto esté compuesto por átomos. Los estados del átomo son sus estados de energía permitidos; las entradas y salidas son fotones, esto es, cuantos de energía electromagnética. En una descripción precisa, me parece que el átomo habría de considerarse como máquina de Mealy no determinística con transiciones épsilon. Es no determinístico, porque resulta imposible la predicción cierta de los efectos de una determinada entrada. Es una máquina de Mealy porque la naturaleza de la salida (a saber, la energía del fotón) está determinada por la transición, y no por el estado al que se da entrada. Las transiciones épsilon son las que pueden darse en ausencia de señal de entrada; es preciso incluirlas en el modelo, porque un átomo puede emitir un fotón y cambiar de estado espontáneamente.

¿Será descriptible el cerebro como máquina con finitud de estados? Se da la circunstancia de que el estudio de los sistemas estado-finitos comenzó con un modelo de redes neuronales, introducido por Warren S. McCulloch y Walter Pitts en 1943. Las neuronas de McCulloch y Pitts eran células simples con entradas excitadoras e inhibidoras; cada célula tenía una única salida y dos estados internos: de descarga y de reposo. Las células podían organizarse en redes capaces de efectuar diversas funciones lógicas, incluidas las operaciones "o", "y", y "no" que son hoy elementos normalizados de los sistemas lógico-electrónicos. La equivalencia de las redes neuronales idealizadas y de los diagramas de transición entre estados, al estilo de los aquí mostrados, fue establecida en 1956 por Stephen C. Kleene, de la Universidad de Wisconsin en Madison.

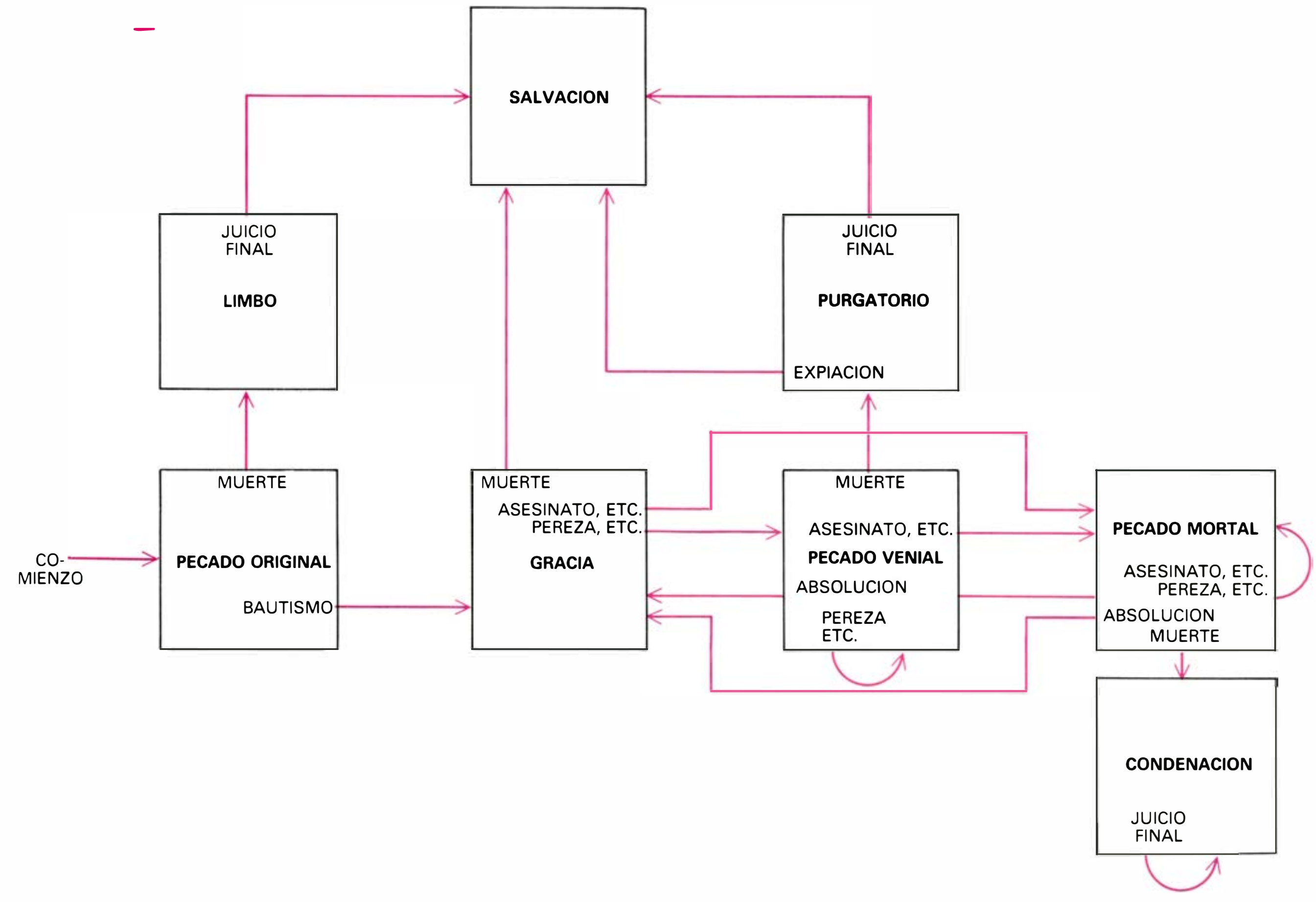
Cuarenta años han pasado desde los trabajos de McCulloch y Pitts, y sigue siendo materia de controversia si es posible clasificar razonablemente al cerebro como sistema de número finito de estados. Evidentemente, el número de neuronas es necesariamente finito,

pero no es éste el meollo de la cuestión. Una neurona auténtica es mucho más compleja que una célula biestable, y muchas de sus propiedades pueden variar sobre una gama continua, en lugar de verse constreñidas a ocupar estados discretos. Además, la prohibición de almacenamiento de información en sistemas auxiliares, característica de los sistemas estado-finitos, resulta, en el mejor de los casos, difícil de adaptar. Si la vida mental no es más que una sucesión de estados instantáneos, carente del conocimiento de la historia propia, ¿qué es entonces la memoria?

Los estados de la mente que la psicología caracteriza, tales como aburrimiento, temor, sed, éxtasis, pena... parecen más fáciles de encajar en una teoría de finitud de estados. Por otra parte, los estados a considerar son tan numerosos, y se comprende tan mal las transiciones entre ellos, que el modelo carece de utilidad. Tan sólo para los animales inferiores es posible trazar algo más que fragmentos aislados del diagrama de transición entre estados, mas, en tales especies, el experimentador puede no tener acceso directo a los estados mentales cuya existencia se presume. Más aún, gran parte de los trabajos de este tipo han sido realizados por conductistas que niegan la existencia misma de estados mentales.

También resulta problemático encuadrar aquí al ordenador digital, y me estoy refiriendo ahora a lo que es la máquina, material y tangible. El modelo mental que se tiene de un ordenador es, por lo común, el formulado por John von Neumann, y divide a la máquina en unidad de proceso y en una formación o matriz de células de memoria. No cabe duda de que la noción de finitud de estados es aplicable a los diversos componentes del procesador central, tales como registros, acumuladores, unidad de control que dirige los procesos internos de la unidad central, etcétera. Las dificultades comienzan tan pronto se toma la memoria en consideración. Con las reglas de construcción de máquinas de tipo finito de estados no es lícita la existencia de una memoria externa y, por consiguiente, cada célula de la memoria del ordenador tiene que considerarse parte del estado global de la máquina, no como un medio de almacenamiento separado y distinto del procesador. Si todas las células de memoria están en blanco, la máquina se encuentra en un estado; si hay una sola ocupada, en otro, y así sucesivamente. Tal concepción del ordenador es particularmente poco esclarecedora, en parte porque no establece conexión entre los estados de la máquina y lo que se encuentra haciendo. Además, el número de estados es inmenso. Ni siquiera un ordenador trivial (100 elementos binarios) que hubiese funcionado desde el nacimiento del universo habría terminado de recorrer todos sus posibles estados.

El papel que desempeña la máquina de número finito de estados está, en principio, a un nivel de abstracción superior al de los mecanismos del soporte físico. Un ordenador, al funcionar bajo la dirección de un programa, deja de ser un ensamblaje de puertas lógicas, registros, células de memoria y demás artilugios de "ferretería" electrónica. Es una máquina "virtual", cuyos elementos funcionales están definidos por el programa, y pueden redefinirse conforme va siendo necesario. Mientras el *hardware* tan sólo entiende de enteros binarios y de mandos simples para desplazarlos y manipularlos, el ordenador virtual maneja sistemas simbólicos mucho más expresivos: palabras, ecuaciones, series, funciones, vectores, codones, listas, imágenes y, quién sabe si, alguna vez, ideas. Las



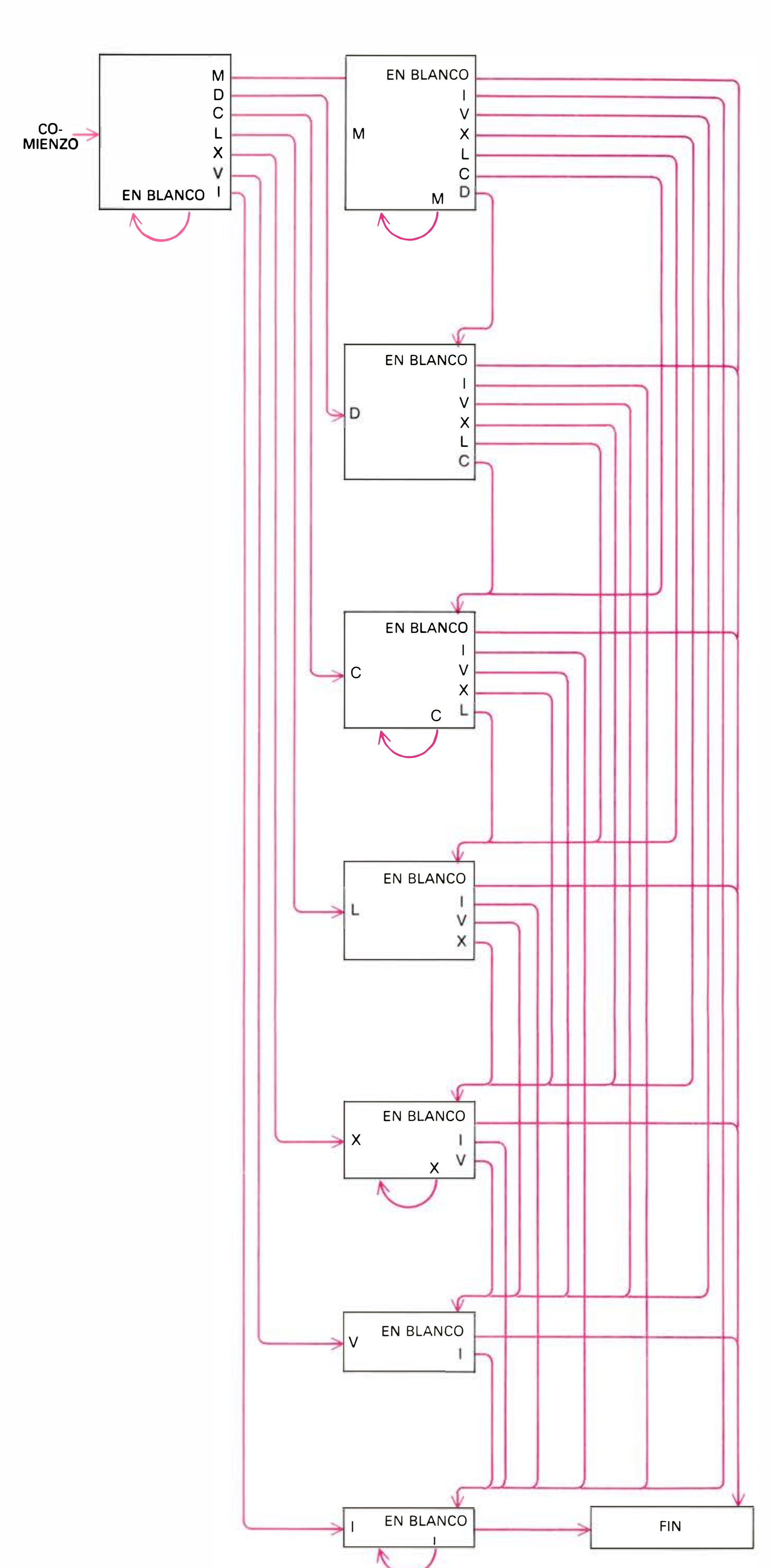
3. Estados del alma, según la teoría de Santo Tomás

técnicas de finitud de estados pueden ser valiosas para crear el ordenador virtual y, en ocasiones, *es* el ordenador virtual una máquina de número finito de estados.

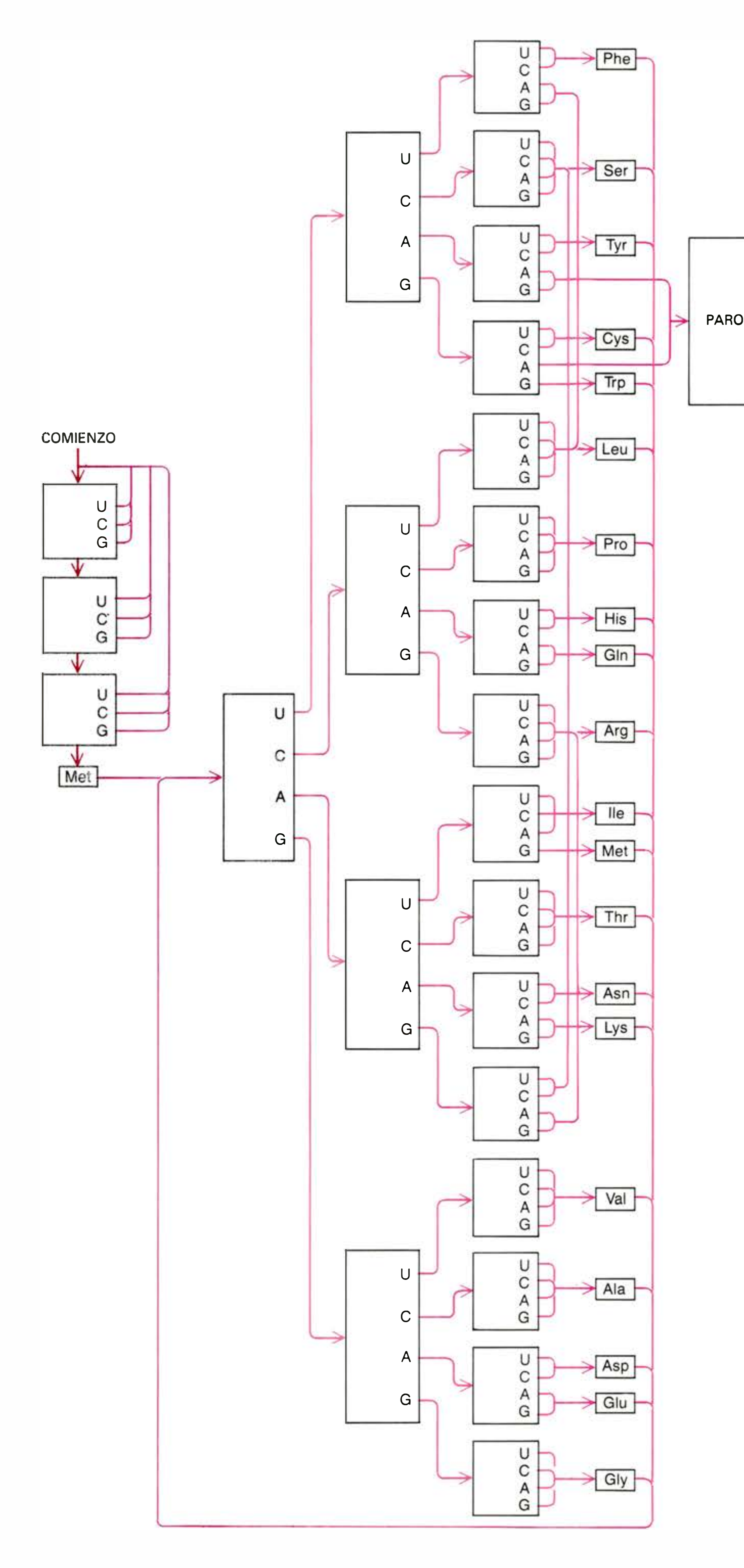
Fijémonos en un programa cuyo objetivo sea leer una serie de dígitos binarios (unos y ceros) e informar de si el número de unos recibido es par o impar. (La tarea reviste importancia práctica; por ejemplo, se utilizan programas de verificación de paridad al transmitir datos telefónicamente.) El programa puede construirse como máquina de tipo finito, con dos estados, como vemos en la figura 2. La máquina comienza a funcionar en estado par, porque hasta ese instante no ha podido recibir ningún 1, y el número 0 se considera par. Cada 1 llegado con el flujo de entrada provoca un cambio de estado, mientras que un 0, recibido en un estado cualquiera, deja invariable tal estado. Aun cuando la máquina no puede "recordar" ninguna de las entradas que precedieron a la más reciente, y no pueda, evidentemente, contar ni los ceros ni los unos, su salida refleja en todo momento la paridad del flujo de ingreso.

El modelo computacional de tipo finito hace su aparición más frecuente en programas que traten de una u otra forma con textos u otras informaciones que adopten formas lingüísticas. Encontramos el ejemplo más preeminente en los compiladores, a saber, programas que traducen los enunciados de programación dados en un lenguaje fuente a enunciados equivalentes en un lenguaje objetivo, por lo general, el "lenguaje de máquina" de algún ordenador concreto. Los compiladores, y otros programas traductores, son esenciales para la noción de máquina virtual; son mediadores entre símbolos con significado para los humanos y los reconocidos por el ordenador.

Todo compilador contiene una parte que puede caracterizarse como máquina de número finito de estados, conocida por analizador de léxico. Lo mismo que el molinete del metro, es un dispositivo traga-fichas. En este caso, empero, las fichas son las palabras, esto es, las unidades léxicas fundamentales del lenguaje. El analizador, mientras examina cada grupo de caracteres, determina si es o no una señal genuina, como por ejemplo lo son números e instrucciones. Si un grupo de caracteres no define una señal reconocible, el analizador lo rechaza, declarándolo absurdo, lo mismo que el molinete rechazaría un pedazo de metal informe.



4. Analizador lexicográfico para un lenguaje de cifras romanas



5. Una máquina estado-finita traduce el código genético a proteína

Podemos ilustrar el funcionamiento de un analizador lexicográfico con una máquina estado-finita diseñada para reconocer los signos de un lenguaje sencillo, si bien, de alcance expresivo muy reducido. Los signos de este lenguaje son, exclusivamente, números romanos. Más todavía, solamente son aceptables cifras romanas de formas especiales, pues es preciso que vengan dadas en notación estrictamente aditiva, con lo que 9 se simboliza por VIIII, y no por IX. (Existen pruebas de que los propios romanos emplearon la notación aditiva; se cree que las formas sustractivas pudieron ser innovaciones de origen germánico.)

Vemos en la figura 4 un diagrama de transiciones entre estados. El alfabeto de símbolos de entrada está formado por las letras *M, D, C, L, X, V* y *I*, amén del espacio en blanco. Cualesquiera espacios en blanco iniciales son despreciados; una vez recibida la primera letra, el programa inicia una transición, saltando a un estado identificado (por comodidad) con el nombre de la letra. Si la primera letra es una *M*, puede ser seguida por cualquier carácter del conjunto permitido, incluida otra *M*. Empero, si el carácter siguiente es una *D*, cambia la situación. Desde el estado *D* no está definida transición alguna al estado *M*, porque ninguna serie de símbolos que contenga *DM* puede ser un signo bien formado en el lenguaje de los números romanos, en expresión aditiva. Además, tampoco hay transición desde el estado *D* al propio estado *D*, así que la sucesión *DD* está excluida. (La razón es que los símbolos de valor intermedio, a saber, *D, L* y *V*, no pueden repetirse en números romanos.)

Una vez en estado *D*, las únicas letras reconocibles son las de valores inferiores, *C, L, X, V* e *I*. Este mismo conjunto es el aceptado en el estado *C* (pues *C* sí puede repetirse); en cambio, en el estado *L* tan sólo son reconocibles las letras *X, V* e *I*. Debería saltar a la vista la regla que gobierna las transiciones. Los estados de la máquina están organizados jerárquicamente y, una vez alcanzado un nivel dado, la máquina no puede retornar a estados superiores; los valores intermedios son "metaestables": la máquina no puede permanecer en ellos. En cuanto se alcanza el estado *I*, sólo es permisible una *I* o un espacio en blanco. El espacio en blanco, se introduzca en este punto o en cualquier otro momento posterior a la entrada de una letra, indica el fin de la señal, y devuelve a la máqina al estado inicial primitivo, presta para recibir el siguiente número romano.

Ninguno de los lenguajes de programación de que tengo noticia permite la introducción de valores numéricos mediante cifras romanas, pero virtualmente todos ellos permiten el manejo de cifras arábigas. Las técnicas para reconocimiento de éstas son parecidas, aunque es mayor su variedad de formatos. Números enteros sencillos, como 137, pueden ser manejados sin dificultad por máquinas de estado único, pero el manejo de las diversas partes de números tales como $+6{,}625 \times 10^{-27}$ requiere un análisis lexicográfico más elaborado.

El sistema ribosoma-ARN de transferencia puede considerarse un analizador de léxico capaz de reconocer secuencias de nucleótidos, biológicamente significativas, en el seno de una molécula de ARN mensajero. Para ser aceptada, la secuencia ha de comenzar por un codón de "puesta en marcha" y finalizar con uno de los tres codones de detención; entre tales límites, cualquier combinación de los símbolos de entrada *U, A, G* y *C*, tomados de tres en tres, está permitida.

El análisis lexicográfico es tan sólo el primer paso en el proceso de compilación. Una vez efectuado, han de entrar en acción otros componentes, llamados analizador sintáctico y generador de código. El analizador sintáctico recibe por su entrada los signos reconocidos válidos en el análisis de léxico; determina sus relaciones sintácticas mutuas. Es lo más cerca que puede llegar a estar el compilador de "comprender" el significado de los enunciados de programa que está traduciendo. El generador de código escribe un programa en el lenguaje objeto, que será el encargado de hacer ejecutar las funciones especificadas por enunciados analizados sintácticamente.

En los lenguajes "de juguete" aquí comentados, los papeles del analizador sintáctico y del generador de código son triviales. La forma compilada de una expresión en números romanos podría ser, sencillamente, la equivalente con cifras arábigas. Que podría ser generada con la siguiente estrategia. Antes de examinar un determinado grupo de signos, se especifica una célula de almacenamiento, y se le asigna el valor numérico 0. A continuación, cada vez que el analizador entre en estado *M* se suman 1000; para el estado *D*, se suman 500, y así sucesivamente. Cuando quede terminado el análisis lexicográfico, la célula de memoria estará cargada con el valor numérico de la cifra romana. Obsérvese que este compilador, aunque de juguete, ya no es una máquina estado-finita pura, porque dispone de un elemento auxiliar de almacenamiento de información.

Aún más sencillo sería un compilador para código genético, que es totalmente realizable en el contexto de los sistemas estado-finitos puros. El "programa" compilado es una sucesión formada por los símbolos de tres letras con que suelen designarse los aminoácidos; tales símbolos pueden ser generados como salida de los estados del analizador que identifican cada codón. Los tres estados correspondientes a codones de interrupción no disponen de salida.

La creación de un compilador adecuado a un lenguaje lo bastante extenso para ser de utilidad general no es tarea trivial; pero, al menos, los principios arquitectónicos subyacentes de construcción de máquinas estado-finitas proporcionan un principio organizador. Si se especifica con precisión suficiente la sintaxis del lenguaje, cabe incluso mecanizar parte del trabajo, que puede ejecutarlo un compilador de compilador, esto es, un programa cuya entrada sea una descripción formalizada de un lenguaje, y cuya salida sea otro programa que traducirá enunciados expresados en dicho lenguaje. Que yo sepa, nadie ha redactado un compilador de compilador de compilador.

El reconocimiento e identificación de los signos que realiza el analizador de léxico es, en sí mismo, una especie de análisis sintáctico; el conjunto de todas las posibles sucesiones de símbolos que pueden formar parte de los signos constituye una especie de lenguaje. Más aún, a menos que arficialmente se imponga alguna clase de limitación, es un lenguaje infinito, pues puede formarse una variedad infinita de signos reconocibles. ¿Cómo podrá una máquina que consta solamente de un número finito de partes reconocer una infinidad de enunciados bien formados, y desechar la infinidad de los incorrectamente construidos? La clave reside en la estructura del lenguaje mismo. Si una máquina con finitud de estados ha de poder reconocer los enunciados de un lenguaje infinito, éstos habrán de construirse de conformidad a reglas estrictas.

Reglas que fueron enunciadas por Kleene en 1956; definen una categoría de lenguajes llamados lenguajes regulares o conjuntos regulares. Kleene demostró que una máquina estado-finita tan sólo podrá reconocer un lenguaje cuando éste sea regular, y más todavía, que todo lenguaje regular es reconocible por alguna máquina estado-finita. El significado de "regular" puede indicarse sucinta (ya que no rigurosamente) mediante dos reglas. Primera, todo lenguaje finito es regular; puede por tanto ser reconocido por una máquina de tipo finito; después de todo, se podría construir una máquina que tuviera un estado para cada posible expresión formada en el lenguaje. En segundo lugar, si un lenguaje es infinito, tienen que poderse analizar todos sus enunciados leyendo uno por uno sus símbolos de izquierda a derecha, es decir, de principio a fin, sin tener que repasar lo leído ni saltar hacia delante. Si la aceptabilidad de un símbolo depende de la presencia de otro símbolo, el símbolo de gobierno ha de encontrarse precediéndolo inmediatamente.

La segunda regla es consecuencia directa de las limitaciones de la máquina de número finito de estados, que no puede prever sus estados futuros ni llevar registro de los pasados; ha de elegir la transición de estados basándose tan sólo en su estado actual y en el símbolo de entrada actual. Por este motivo, la máquina estado-finita es incapaz de manejar para los números romanos la notación sustractiva. Si se lee la expresión XI, y se interpreta como 11, la máquina no puede retroceder y revisar el valor cuando el siguiente carácter resulte ser V. Muchas otras funciones quedan descartadas por esta misma limitación. Por ejemplo, no es posible construir una máquina estado-finita capaz de leer una secuencia de dígitos binarios y determinar si el número de unos es igual al de ceros. Análogamente, aunque una máquina estado-finita sea capaz de sumar números binarios, no puede multiplicarlos. Dejo al cuidado del lector deducir por qué.

Más allá de las máquinas estado-finitas y de los lenguajes regulares se extiende toda una jerarquía de máquinas más potentes y de lenguajes más generales. Es la llamada jerarquía de Chomsky, en honor del lingüista Noam Chomsky, quien investigó los diversos lenguajes formales como posibles modelos del lenguaje natural. Se crean lenguajes más generales relajando las restricciones que constriñen a los conjuntos regulares; las máquinas se construyen añadiendo elementos de memoria al modelo fundamental, la máquina de número finito de estados.

La máquina siguiente de la serie es llamada autómata descendente. Consta de una máquina de tipo finito con la adición de una estructura de memoria, de capacidad infinita, pero de organización peculiar. La memoria adopta la forma de una "pila", a modo de rimero

bien equilibrado de bandejas en un autoservicio. La única forma de almacenar un nuevo elemento de información es situándolo en lo alto de la pila; cuando se precisa recobrar una información hay que retirar primero cualesquiera elementos apilados encima. El último elemento ingresado en la pila será el primero en salir.

Los lenguajes reconocibles por los autómatas descendentes son llamados lenguajes insensibles al contexto o, brevemente, lenguajes sin contexto. Al analizar sintácticamente sus enunciados, la aceptabilidad de un símbolo dependerá tanto del símbolo situado a su izquierda como de su contiguo derecho inmediato. Esta dependencia bidireccional es permisible porque cualesquiera símbolos cuya interpretación no pueda quedar decidida en el acto quedarán almacenados en la pila hasta que se resuelva la ambigüedad. Por tanto, un autómata descendente puede funcionar con números romanos en notación sustractiva, y reconocer e identificar expresiones con iguales números de ceros y unos (u otros símbolos, como apertura de paréntesis y cierre del mismo). Por otra parte, es incapaz de detectar enunciados que contengan iguales números de tres símbolos, como ceros, unos y doses. La mayoría de los lenguajes de programación son insensibles al contexto; por lo común, el analizador sintáctico de los compiladores es un autómata descendente. En uno de los lenguajes de programación, el llamado Forth, la estructura primaria de la memoria es una pila. Evidentemente, en ninguna máquina real podrá tener la pila profundidad infinita.

Los lenguajes insensibles al contexto merecen tal nombre porque el análisis de un símbolo cualquiera solamente puede quedar influido por los dos símbolos vecinos inmediatos del dado. Al examinar esta restricción aparecen lenguajes "sensibles" al contexto, al tiempo que aumenta la dificultad de interpretación. Ahora pueden interactuar símbolos muy separados. En el peor de los casos, puede ser imposible interpretar el primer símbolo de una expresión hasta no haber leído el último. A cambio de las dificultades adicionales así creadas, se gana algo en la capacidad de tratamiento. Las máquinas basadas en lenguajes sensibles al contexto pueden determinar si una expresión contiene iguales números de tres símbolos.

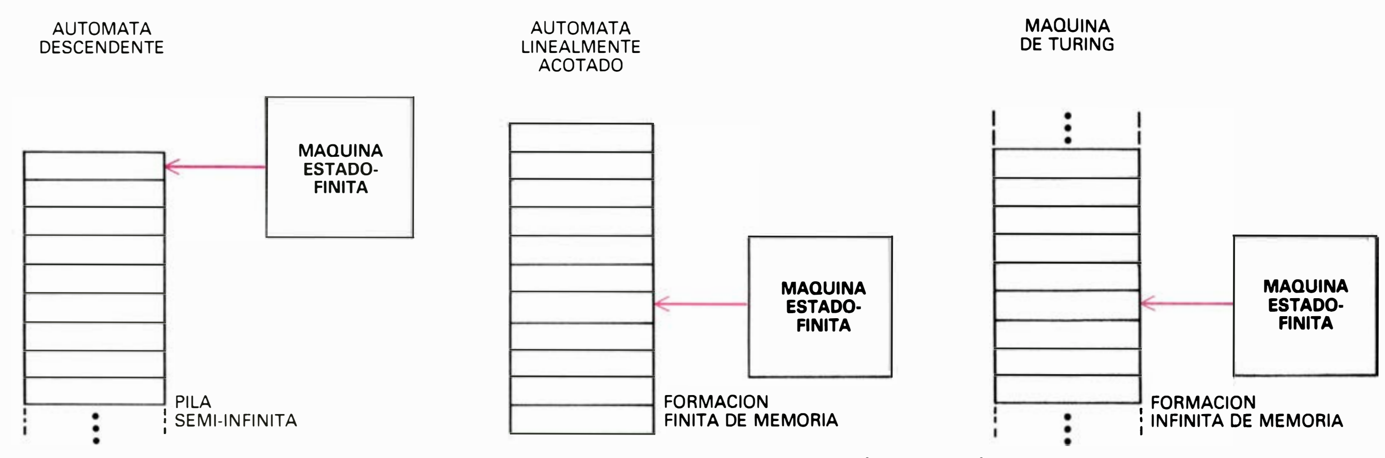
La máquina capaz de reconocer lenguajes sensibles al contexto es el autómata linealmente acotado. Además del aparato de número finito de estados ordinario, dispone de una memoria organizada de modo tal que cualquiera de sus alojamientos de memoria es alcanzable en cualquier momento; es una máquina de acceso aleatorio. La memoria es de capacidad finita, pero se la supone suficientemente grande para admitir cualquier entrada que la máquina reciba. El autómata linealmente acotado parece una buena aproximación del modelo de von Neumann de ordenador digital. Curiosamente, no abundan los lenguajes de programación sensibles al contexto; al parecer, la estructura insensible al contexto, más sencilla, tiene casi siempre suficiente capacidad expresiva.

Todos los lenguajes descritos hasta ahora comparten una propiedad: son lo que se llama recursivos. Lo que esta designación quiere expresar es que se puede imaginar un procedimiento para generar todas las posibles "manifestaciones" del lenguaje, en orden de longitud creciente. Se deduce que existe un método garantizado para decidir si un cierto enunciado de longitud finita pertenece o no al lenguaje: es suficiente generar todos los enunciados hasta la longitud correspondiente y compararlos.

Existen lenguajes que no pueden siquiera cumplir esta norma mínima de tratabilidad. Tan sólo hay para ellos una posible máquina capaz de reconocerlos: el ordenador definitivo, la máquina de Turing, un autómata estado-finito al que se permite retozar libremente por una memoria ilimitada. En la descripción dada por Turing, la memoria es una cinta dividida en casillas, ilimitada en ambas direcciones, sobre la cual un aparato estado-finito puede escribir, leer o borrar lo escrito.

Desde la elevada perspectiva de la máquina de Turing, se aprecian más claramente las relaciones de los dispositivos de cómputo de rango inferior. El autómata linealmente acotado es, en puridad, una máquina de Turing dotada de una cinta finita. El autómata descendente tiene una cinta ilimitada en una dirección, pero en ella la "cabeza" de lectura y grabación sobre la cinta permanece siempre fija sobre la última casilla que no está en blanco. La máquina de número finito de estados es una máquina de Turing totalmente desprovista de cinta.

Lectores habrá que, ansiosos por comenzar el análisis sintáctico de lenguajes no recursivos, salgan de compras dispuestos a adquirir una máquina de Turing. Es preciso advertirles que incluso el ordenador definitivo tiene su talón de Aquiles. Existen lenguajes de gramática tan absurda que ni siquiera con una máquina de Turing podemos estar seguros de reconocer sus enunciados en un tiempo finito. Hasta el momento, tales lenguajes han mostrado poca utilidad en el mundo de las máquinas de cómputo. No se sabe muy bien cómo, pero la gente se las arregla para hablarlos.



6. La jerarquía de Chomsky entre máquinas finitas e infinitas

Taller y laboratorio

Cuando vertimos gotas de aceite en agua suceden cosas muy curiosas

Jearl Walker

¿Qué acontece cuando se vierte una gota de aceite, o de algún otro líquido, en un recipiente con agua? Si la gota es soluble en agua, desaparecerá rápidamente en la masa de ésta. Si es insoluble, piénsese en el aceite de parafina, permanecerá donde la hayamos dejado. Si su solubilidad es intermedia, se extenderá por la superficie, hasta convertirse en una capa delgada e invisible; si añadimos luego más gotas, formarán manchas o charcos. ¿Qué es lo que determina el comportamiento de la gota?

La respuesta a esa pregunta permanece, hasta cierto punto, en el aire. Sabemos que el comportamiento de una gota depositada sobre la superficie del agua lo determinan las complejas interacciones (en su mayoría eléctricas) de las moléculas de la superficie de separación. Recientemente, Frode Wissing, del Real Instituto Dental de Dinamarca, me ha remitido un manuscrito de sus experimentos en torno a la solubilidad en agua del ácido oleico (uno de los principales componentes del aceite de oliva) y del aceite de parafina. Este trabajo, ideado como experiencia de cátedra, descubre algunas curiosas piezas del enigma.

La solubilidad en agua de un compuesto químico depende de la importancia del enlace entre las moléculas del mismo y las del agua, de tal modo que el grado de solubilidad es resultado de una pugna entre los enlaces que mantienen unida una molécula y las demás posibilidades de enlace que ofrece la otra sustancia.

La solubilidad en agua de los compuestos orgánicos varía entre márgenes muy amplios. Sin esta variabilidad, la vida no sería posible en la Tierra. Los compuestos orgánicos insolubles constan de grupos de átomos que se enlazan poco (y, a veces, nada) con las moléculas de agua. De estos grupos se dice que son hidrófobos, al igual que de las moléculas que sean insolubles por su causa. Este adjetivo, engañoso, parece insinuar que exista una repulsión entre la molécula (o el grupo) y el agua. Pero no hay tal, ya que el efecto tiene su origen en la debilidad del enlace, tanta que la cohesión del agua excluye el compuesto hidrófobo.

Un ejemplo de compuesto hidrófobo lo tenemos en el aceite de parafina, formado exclusivamente por grupos hidrófobos de átomos de carbono e hidrógeno. Estos grupos hidrocarbonados se enlazan poco, o nada, con las moléculas de agua; toda gota de aceite de parafina que se deposite sobre la superficie del agua permanecerá, pues, intacta. Por su peso, la gota empujará contra la superficie del agua, pero no se disolverá en ella y ni siquiera se esparcirá.

Muchas otras moléculas orgánicas son parcialmente solubles, porque al menos alguno de sus grupos atómicos se enlaza con el agua. Entonces, cuanto mayor número de tales grupos tenga el compuesto, más soluble será. Por ejemplo, la glucosa es un azúcar que se disuelve en agua sin dificultad, porque posee seis grupos hidróxilo (OH) que se enlazan con las moléculas de esta última.

El ácido oleico se encuentra en una posición intermedia entre la parafina y la glucosa. En un extremo de su molécula tiene un grupo COOH, que se enlaza con el agua. El resto de la molécula, constituido por una cadena de grupos hidrocarbonados, es hidrófoba. Esta doble naturaleza de la molécula del ácido oleico responde de su capacidad para extenderse por la superficie del agua hasta formar una capa de una sola molécula de espesor.

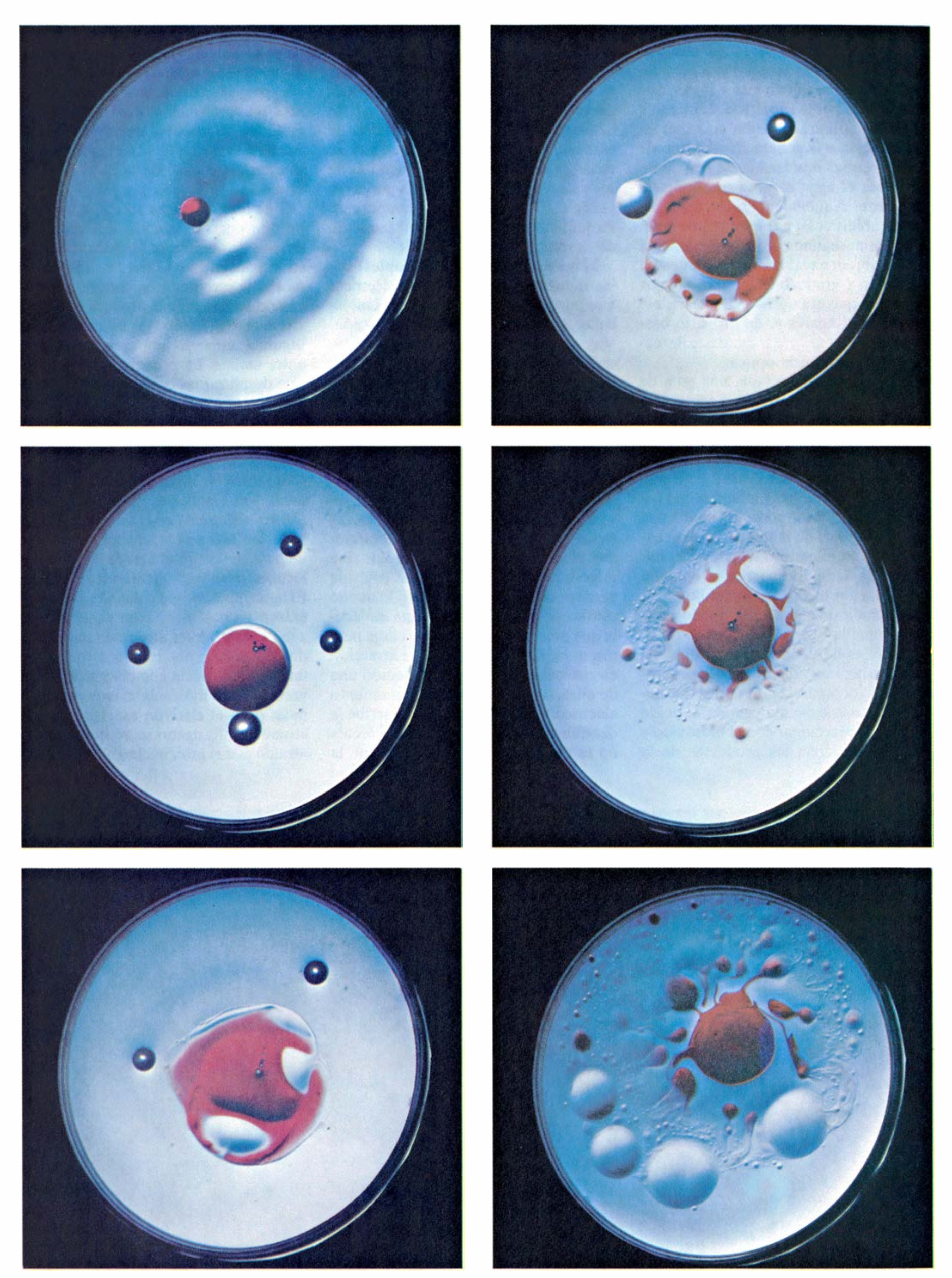
Wissing se proponía ilustrar las diferencias entre las solubilidades del aceite de parafina (estrictamente hidrófobo) y del ácido oleico (parcialmente hidrófilo). Sus experiencias las llevó a cabo en platillos de Petri de nueve centímetros de diámetro perfectamente limpios. (Cualquier contaminación, aun la de una huella digital, deposita una monocapa en la superficie del agua y puede alterar los resultados.) En un proyector suspendido del techo, que rodeó de papel oscuro para que la luz pasara nada más que por el platillo, instaló uno de éstos. Con semejante dispositivo consiguió ampliar diez veces la imagen de lo que ocurría en el platillo, algo más que suficiente para que los alumnos siguieran lo que acontecía en el platillo cuando él agregaba gotas de fluidos orgánicos. Estas experiencias las he repetido a plena satisfacción en mi cocina, aunque empleando aceite de oliva en vez de ácido oleico.

Wissing comenzó poniendo de manifiesto la insolubilidad del aceite de parafina depositando una gota del mismo sobre agua destilada. El aceite, insoluble, se limitó a quedarse encima del agua; adquirió una forma de lente biconvexa. Cuando añadió más gotas, algunas se fusionaron para formar una lente mayor, pero ninguna se extendió por la superficie ni se disolvió.

Realizó luego tres ensayos similares con platillos llenos de otras soluciones: ácido clorhídrico, hidróxido sódico e hidróxido amónico, todos con una concentración 0,1 molar. El aceite formó siempre una lente, sin disolverse en la solución ni esparcirse sobre ella.

Al repetir el experimento con una gota de ácido oleico, consiguió un resultado más espectacular. Había teñido el ácido oleico con un poco de rojo Sudan III, para así observarlo mejor. Cuando puso una gota del ácido sobre el agua pura, en la superficie de ésta se generó una onda de choque, como si alguien hubiera golpeado el platillo. La gota se extendió entonces con toda rapidez por la superficie, dejando una película tan tenue que ni siquiera se distinguía el tinte rojo. Además, parte del ácido formó una gota biconvexa de gran tamaño y otras más pequeñas.

En mis ensayos desplegué una gota de aceite de oliva diminuta desprendiéndola, de una sacudida, de un clip sujetapapeles desdoblado y endereza-



1. Experimentos de Frode Wissing en los que una gota de ácido oleico engulle gotas de parafina vertidas en agua

do. (El cual había previamente limpiado con detergente y luego calentado al rojo uno de sus extremos.) Esa gota desapareció inmediatamente a la vez que se extendía por la superficie; formó una capa cuya finura impedía su observación directa o la generación del tipo de interferencias ópticas que se ven cuando sobre el agua se posa una tongada de aceite. Aquella superficie debía de estar cubierta casi por entero de aceite, pues una segunda gota se mantenía unida durante varios minutos antes de extenderse gradualmente para formar una capa que era perfectamente visible en algunos lugares y, en otros, lo bastante gruesa para crear bandas coloreadas de interferencia óptica.

Cuando Wissing añadía una gota de ácido oleico a un platillo con ácido clorhídrico diluido, también se extendía por toda la superficie para desarrollar una gota grande. Pero entonces la lente era más plana y llenaba casi todo el platillo. El ácido oleico añadido a una solución de hidróxido sódico formaba una lente irregular, de la que escapaba suavemente en hilos; transcurridos 20 minutos, la lente desaparecía, dejando una solución turbia en el platillo.

Al añadir una gota al hidróxido amónico, aquélla se transformaba en hilos, todos ellos bordeados de zonas claras, que corrían rápidamente por la superficie para desaparecer después. La superficie se encalmaba cuando el ácido cesaba de reaccionar con el hidróxido y en ella quedaban los productos de la reacción, aunque la masa líquida no mostraba turbidez alguna.

Wissing se fijó la tarea de explicar tales variaciones en función de las fuerzas actuantes entre el ácido oleico y las distintas soluciones. La más sencilla de esas fuerzas es la gravedad, que atrae a cada gota de ácido hacia el seno de la solución y tiende a extenderla por la superficie. Otra fuerza es el empuje de Arquímedes, debido al desplazamiento de la solución.

Más complicadas (e interesantes) son las numerosas interacciones eléctricas que pueden tener lugar entre las moléculas de la gota y las de la solución. Voy a exponerles primero los modelos sencillos que la mecánica cuántica adopta para estas fuerzas y luego la aplicación que de ellos hace Wissing a sus experimentos. Téngase presente, empero, que esos modelos se resienten de la falta de conocimientos más detallados de mecánica cuántica.

Cuando una molécula dispone de lugares con una carga eléctrica neta, es capaz de enlazarse merced a interacciones electrostáticas. Por ejemplo, una molécula que tenga un punto con una carga neta positiva atraerá, en lo que se califica de enlace iónico, a una molécula que presente un punto con carga neta negativa. Un segundo tipo de atracción electrostática se desarrolla cuando una de ellas, o las dos, tengan un momento eléctrico. (Término éste que describe la distribución de carga en una molécula no esférica.) En cuyo caso, aunque la molécula no posea carga neta, el centro de su carga negativa (de los electrones) no coincide con el de su carga positiva (de los protones). En tal situación aparece un campo eléctrico en torno a la molécula. Entonces, cuando dos moléculas de este tipo se sitúan cada una en el campo eléctrico de la otra, resultan atraídas.

Una tercera fuerza atractiva también posible aparece cuando una molécula dotada de un momento eléctrico intenso se encuentra en la vecindad de otra molécula carente de él. En este caso, el campo eléctrico de la primera molécula altera la distribución de las cargas de la segunda, creando una mutua atracción entre ambas. La primera molécula podría desplazar las órbitas de los electrones de la segunda, de manera que el centro de la carga negativa se coloque entonces más lejos que el centro de la carga positiva.

Una acción intermolecular prácticamente omnipresente cuando hay momentos eléctricos es el enlace hidrogénico. De éste es responsable un átomo de hidrógeno situado entre las dos moléculas; sirve para mantenerlas unidas. El agua mantiene su cohesión gracias al enlace hidrogénico. Una molécula de agua consta de un átomo de oxígeno y dos de hidrógeno, reunidos formando una *V* muy abierta con el oxígeno en el vértice. Esta molécula carece de carga neta, pero el electrón asociado a cada átomo de hidrógeno se ve fuertemente atraído por el gran núcleo del átomo de





2. Efecto del hidróxido amónico sobre parafina y ácido oleico

oxígeno. De este modo, la molécula posee una distribución interna de cargas que crea un campo eléctrico a su alrededor.

En una descripción más simple decimos que los extremos hidrogénicos de la *V* se hacen positivos y el vértice, donde está el oxígeno, se hace negativo. Pero hay que rehuir de la hipersimplificación que supone afirmar que los átomos de hidrógeno quedan desposeídos de sus electrones; las cosas no son tan radicales. Para evidenciar la presencia de un momento eléctrico, se dice que la molécula de agua es polar.

Imaginemos dos moléculas de agua contiguas alineadas tal como se muestra en la figura 5. La molécula de la izquierda presenta uno de sus átomos de hidrógeno al de oxígeno de la molécula de la derecha, formando así un enlace hidrogénico. Se trata de una acción eléctrica en la que la punta positiva de una molécula es atraída hacia el vértice negativo de la otra. La molécula de la izquierda (que se llama dador de hidrógeno) no pierde su átomo de hidrógeno, ni lo gana la de la derecha (aceptor de hidrógeno). Este hidrógeno intermedio se halla hacia un tercio de la distancia entre un oxígeno y el siguiente; de él se dice que lo comparten las dos moléculas, pese a que el dador siga sujetándolo con más fuerza.

Aunque esta imagen de enlace hidrogénico sirva perfectamente, debo recalcar que no se ha ideado ningún modelo matemático que se haya mostrado plenamente acertado para explicar este enlace. La acción entre dos moléculas de agua no responde fielmente al efecto estático sencillo, aquí descrito. Dentro de la masa de agua, el ordenamiento de las moléculas más estable posible se da cuando todas poseen enlaces de hidrógeno con cuatro moléculas contiguas. Para dos de éstas, la molécula central actúa de dador de hidrógeno; para las otras dos actúa de aceptor. Si bien esta disposición es la más estable de todas, cabe que los enlaces se estiren, giren y se rompan. Sin esta flexibilidad el agua no fluiría. Y más aún, el ordenamiento de las moléculas de agua no es estático, ni siquiera cuando ésta reposa tranquila en un recipiente. Los enlaces se rompen y rehacen continuamente.

Es corriente que un compuesto se disuelva en el agua porque la molécula de ésta sea suficientemente pequeña para acercarse a las moléculas del primero. Si éste dispone de lugares con carga neta, el agua formará enlaces iónicos por atracción electrostática al exponer su átomo de oxígeno o uno de hidrógeno al lugar cargado. En los sitios de la molécula donde haya grupos polares, el agua formará enlaces de hidrógeno con el compuesto.

En general, en un enlace de hidrógeno pueden sólo participar átomos de oxígeno y nitrógeno. Los de carbono no pueden hacerlo porque propenden a rodearse de hidrógeno (para formar grupos hidrocarbonados) y no exponen puntos a los que puedan enlazarse moléculas de agua. Además, no crean momento eléctrico capaz de formar un grupo polar. Esta diferencia entre las aptitudes de enlace del carbono y el oxígeno constituye la razón fundamental por la que los grupos hidrocarbonados del ácido oleico son hidrófobos, mientras que el grupo COOH del extremo de su molécula es hidrófilo.

Otra posibilidad de enlace (aunque escasa) la proporciona la disociación natural del agua en iones H^+ y OH^-. A éstos los atrae el compuesto hacia sus lugares dotados de carga neta o momento eléctrico. Por último, podría ser que el compuesto se disocie aportando un átomo de hidrógeno para una molécula de agua. Entonces el punto negativo que quede en el compuesto atraerá al ion de agua positivo.

Cuando dos moléculas neutras y esféricas estén cerca una de la otra, no debe haber interacción eléctrica. Por tratarse de esferas, el centro de la carga negativa de cada una coincide con el de la positiva y, por ello, carecen de momentos eléctricos. Tampoco hay atracción iónica. Aún así, pueden atraerse merced a una curiosa circunstancia.

En una descripción mecánico-cuántica simple, las órbitas de las moléculas se sincronizan y crean momentos eléctricos instantáneos que originan campos eléctricos fugaces. Estos campos provocan una atracción entre las moléculas, denominada fuerza de van der Waals.

Para describir de modo elemental la fuerza de van der Waals se comienza por considerar los electrones que orbitan en moléculas contiguas. Estas moléculas "ven" una a la otra como eléctricamente neutra, pues las cargas positivas y negativas de su interior son iguales. Sin embargo, cuando las moléculas se acercan mutuamente, los dos grupos de electrones interactúan y desarrollan órbitas síncronas.

Así, podemos suponer que, en un instante dado, una molécula sufra una separación momentánea entre sus cargas que la dote de momento eléctrico. Y lo mismo le ocurrirá a la otra molécula. En dicho instante, ambas se atraerán. Atracción que, aunque fluctúe en el tiempo, lo hará con un valor medio no nulo. La verdad es que es lo bastante intensa para mantener unidas a las moléculas. El aceite de parafina se mantiene cohesionado por fuerzas de van der Waals.

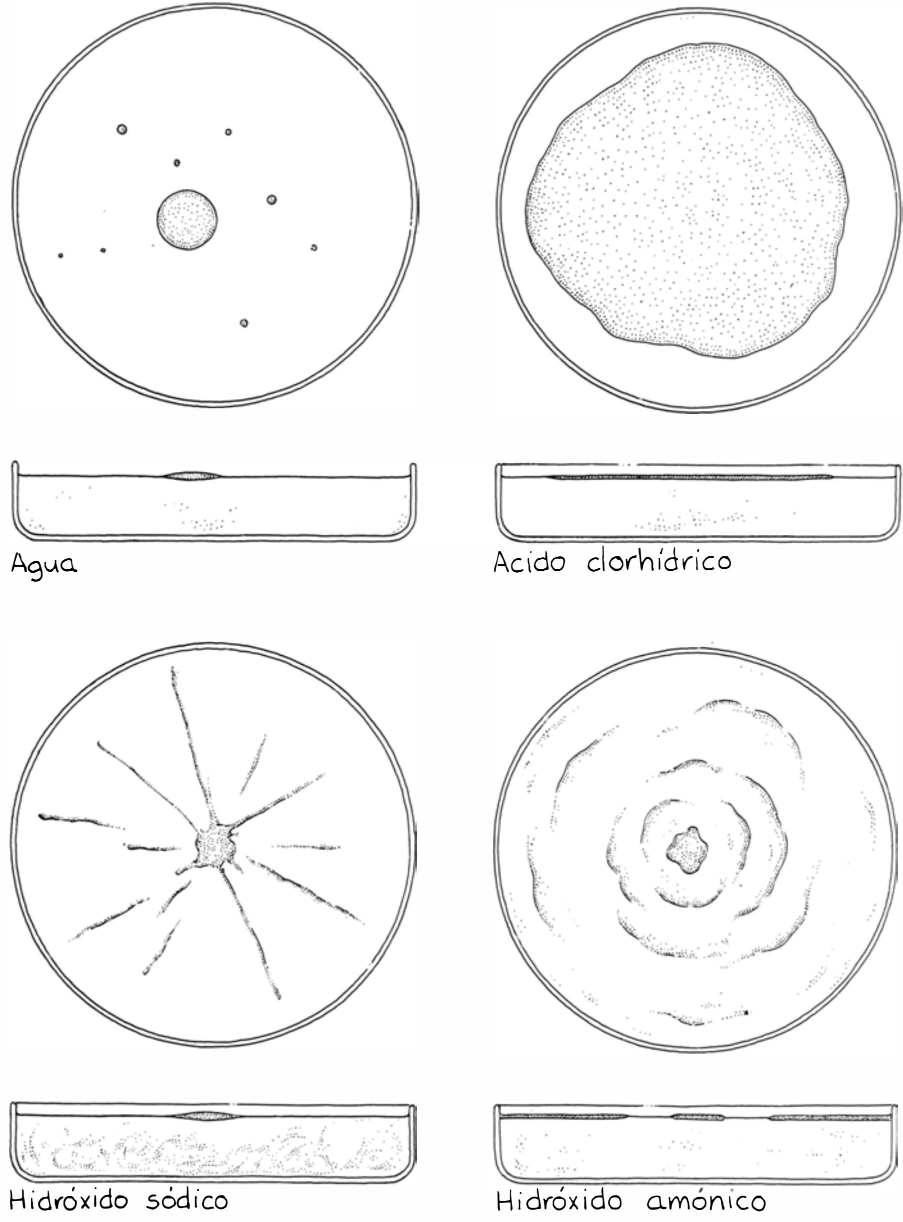
Provisto de este ramillete de modelos de fuerzas intermoleculares, Wissing justificaba sus experimentos y proponía otros nuevos. Al depositar sobre agua una gota de aceite de parafina, no se establece, esencialmente, enlace alguno entre los grupos hidrocarbonados de la molécula de parafina y las moléculas de agua. Entre ambos tipos de moléculas resulta inviable el enlace hidro-

3. Algunas estructuras moleculares fundamentales

génico. Los demás tipos de enlace, rarísimos, no rivalizan con el enlace hidrogénico, que mantiene la cohesión del agua, ni con la fuerza de van der Waals, que mantiene la de la gota de aceite. Por cuya razón, ésta se ve únicamente sometida a la gravedad que la oprime de un modo suave contra la supeficie. Si ponemos la gota en las otras soluciones ocurrirá lo mismo, ya que las posibilidades de enlace en la superficie de separación seguirán siendo escasas.

Con esto se aclara más el trasfondo del comportamiento del ácido oleico. Cuando depositamos una gota de éste sobre agua pura, en la superficie de separación se inicia inmediatamente el enlace. El extremo hidrófilo del ácido es atraído hacia el agua, pero no lo son el extremo opuesto y gran parte de la longitud de la molécula. Corrientemente, el extremo hidrófilo suele recibir el nombre de cabeza de la molécula y el extremo opuesto el de cola. Las moléculas de ácido oleico situadas en la superficie de separación rotan de tal modo que las cabezas se alinean hacia la superficie de agua subyacente, quedando las colas hacia arriba.

Entre la cabeza y el agua aparecen dos tipos de enlace. Como tanto la cabeza como la molécula de agua son polares, sus momentos eléctricos acaban adoptando una configuración que facilita la atracción. (Hablando con mayor propiedad, esta atracción debería describirse como un puente o enlace de hidrógeno. Efectivamente, cada molécula de agua presenta uno de sus átomos de hidrógeno al átomo de oxígeno que expone la cabeza y así forman el enlace.)



4. Acción de gotas de ácido oleico en cuatro soluciones

En el otro tipo de enlace interviene la disociación del grupo de la cabeza, el cual cede un ion hidrógeno (un protón) a una molécula de agua. Esta pérdida deja la cabeza negativamente cargada (COO^-) y al ion de agua recién formado con una carga positiva (H_3O^+). Dadas sus cargas eléctricas contrarias, ambos se atraerán por enlace iónico.

Entonces, en toda la superficie de separación comienzan a formarse otros enlaces adicionales. Cuando las moléculas del ácido oleico se orientan con las cabezas hacia el agua, se unen entre sí ordenadamente. Las cabezas contiguas forman un enlace hidrogénico, de manera tal que el OH de cada cabeza se alinea con el oxígeno de una cabeza vecina. Entre los grupos hidrocarbonados de la cadena que componen el resto de la molécula se establecen más enlaces. Y, en algunos lugares, los tramos de grupos hidrocarbonados contiguos se atraen entre sí mediante fuerzas de van der Waals.

A medida que las moléculas de ácido oleico rotan para ponerse en formación, el ácido que se encuentra en los límites de la gota va estableciendo su enlace con el agua. Así, merced a la fuerza de van der Waals, ese movimiento atrae más ácido oleico hacia el agua, con la que se enlaza y, con ello, arrastra más moléculas procedentes de la masa de la gota. Esta se adelgaza conforme fluye líquido hacia el borde, hasta que no tarda en convertirse en una película de una molécula de grosor. La alineación entre las moléculas contiguas impide que esa película no se expanda desde ese estado.

Cuando se deposita una pequeña cantidad de ácido oleico sobre agua, la monocapa cubre una porción de la superficie de aquélla. Si hay ácido en exceso, aparece la monocapa y el sobrante queda formando una gota (o varias) de geometría biconvexa en la superficie del agua. Una vez cubierta ésta de moléculas de ácido oleico, las demás ya no se enlazan con el agua, sino que se quedan sólo con las fuerzas de van der Waals y las otras interacciones eléctricas que las hacen cohesivas. O sea, las moléculas excedentes se enlazan unas a otras para formar las gotas sobrantes.

Cuando la gota de ácido oleico se deposita sobre ácido clorhídrico diluido el proceso se repite, menos en un detalle. Esta vez la abundancia de iones hidrógeno y cloro en la masa de líquido imposibilita toda disociación sistemática

en las cabezas de las moléculas de ácido oleico. Las moléculas giran también aquí para alinearse con las cabezas hacia abajo y las colas hacia arriba, aunque en este caso los enlaces con el agua sean casi todos hidrogénicos. Y la gota de ácido oleico se extiende, asimismo, por la superficie.

Las cosas cambian cuando se deposita una gota de ácido oleico sobre una solución diluida de hidróxido sódico. Este neutraliza al ácido, dando oleato sódico. En la concentración que usa Wissing en sus ensayos, las moléculas de oleato se reúnen y constituyen unas gotas minúsculas que reciben el nombre de micelas. Las moléculas de la superficie de separación entre una micela y el agua circundante disponen sus cabezas hacia fuera y sus colas hacia dentro. Las cabezas se enlazan hidrogénicamente con el agua y, a veces, mediante una atracción electrostática, la cual aparece cuando un grupo de cabeza cede un ion hidrógeno a una molécula de agua.

En el experimento de Wissing se observa con toda nitidez la formación de micelas. Tras caer sobre la superficie una gota de ácido oleico, manan de ésta hilos de corriente superficiales que desaparecen gradualmente. Bajo esos hilos y la gota, se forman las micelas. Y a medida que éstas van obstruyendo el paso de la luz a través de la solución enturbian la masa de ésta.

El hidróxido amónico neutraliza también al ácido oleico, pero sin formación de micelas. Conforme se va neutralizando el ácido sale enérgicamente despedido de la gota y forma hilos; éstos y la gota se mueven rápidamente por la superficie. Desconozco el por qué de esa expulsión del ácido.

La siguiente serie de experimentos de Wissing la encuentro fascinante. Comienza por agregar tres o cuatro gotas de ácido oleico al agua pura. Volvemos a tener una monocapa y el ácido sobrante queda en forma de una gota de gran tamaño. En la fotografía de la parte superior izquierda de la figura 1 vemos la gota descansando encima del agua, tremolando aún a causa de la brusca formación de la monocapa.

Lejos de la gota de ácido oleico, Wissing deposita cinco de aceite de parafina. Al agregar la primera de éstas, la gota de ácido oleico adquiere una gran extensión. Depositada la quinta gota de aceite de parafina en el platillo, la de ácido oleico emite una corriente dirigida hacia la más cercana de las de parafina y comienza a consumirla. Entonces, durante varios minutos, el ácido oleico se desliza hacia la gota de aceite, la escala y acaba disolviéndose en ella, tiñendo de rosa el aceite. En ese momento la gota de ácido oleico es más gruesa y extensa.

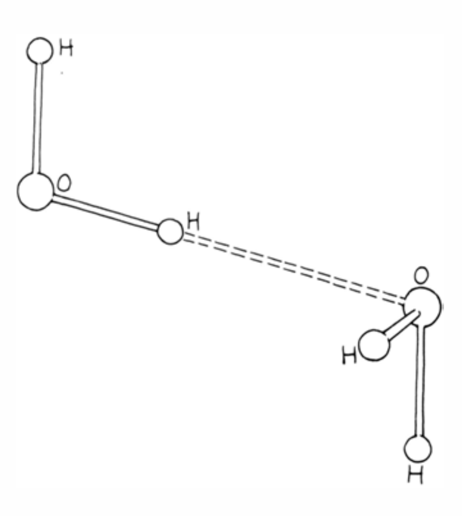
La fotografía de la parte central izquierda de la figura 1 muestra una segunda gota de aceite de parafina inmediatamente antes de ser consumida por la de ácido oleico. En ella se aprecian todavía trazas de la primera gota (ya consumida) a un lado de la gota de ácido oleico. La segunda fotografía, en la parte inferior izquierda, fue tomada después de que la gota de ácido oleico hubiera ingerido una segunda gota y una tercera. El resto del ensayo prosigue en las fotografías que figuran de arriba abajo en la parte derecha.

Ni Wissing ni yo conocemos la razón del ensayo anterior. En su opinión, la presencia del aceite de parafina ofrece un medio merced al cual los grupos hidrocarbonados de las moléculas sobrantes de ácido oleico pueden enlazarse con los grupos hidrocarbonados del aceite. Mi parecer es que, inicialmente, las moléculas de la gota de aceite tiran fuertemente de las colas de las moléculas de ácido oleico existentes en la monocapa. Como éstas se mantienen unidas, el tirón se transmite a la gota de ácido oleico, resultando así que parte de la gota es atraída hacia la gota de aceite.

Este experimento me desconcierta mucho. ¿Por qué una gota de aceite hace que la gota de ácido oleico se extienda, mientras que cinco gotas hacen que el ácido oleico comience a engullir la parafina? Creo que ello tiene que ver con la fuerza horizontal que las gotas de parafina aplican a la monocapa. Supongamos que, antes de depositar la parafina, el platillo esté cubierto sólo por la monocapa y la gota de ácido oleico en exceso. Como el líquido de ésta tiende a resbalar por sus costados, ejerce una fuerza sobre la monocapa.

Cuando se introduce la primera gota de aceite, se crea más fuerza sobre la monocapa. Y cuatro gotas más intensifican la fuerza. Por entonces, una gota de aceite habrá comenzado a enlazarse con las moléculas de ácido oleico que la rodeen.

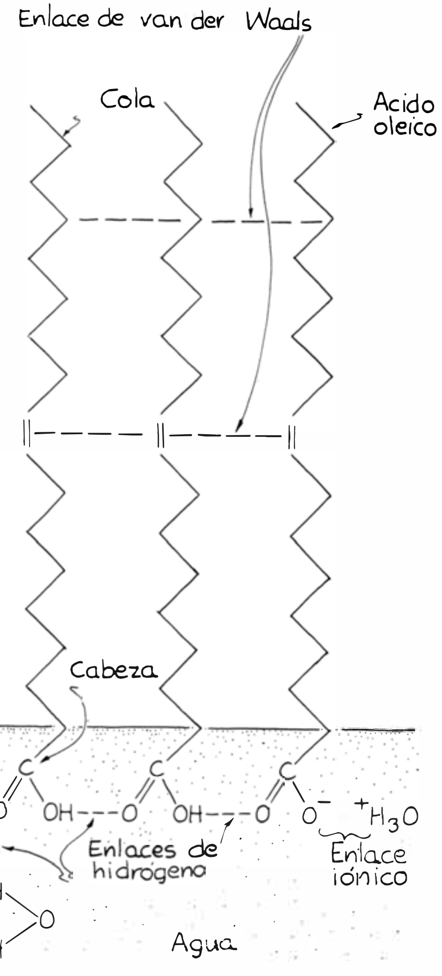
¿De qué dirección procederá la masa renovada de ácido oleico? Preferentemente de la correspondiente a la gota del mismo, a causa de la fuerza con la que actúa contra la monocapa. Así, la gota de ácido se dispersa sobre la monocapa. Esta extensión la estira la monocapa (por culpa del enlace con la gota de aceite) y la comprime la presión



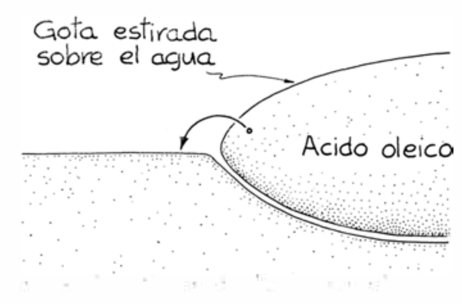
5. Enlace de hidrógeno en el agua

hidrostática debida al fluido del resto de la gota de ácido. Finalmente, la expansión alcanza y devora la gota de aceite.

Wissing cambia a veces la marcha del experimento cuando el ácido oleico está engullendo las gotas de aceite. Para ello vierte uno o dos mililitros de hidróxido amónico concentrado junto a la gota de ácido oleico. Entonces, ésta



6. Monocapa de ácido oleico



7. Así se forma una monocapa

se pone a bailar frenéticamente, a la vez que el nuevo compuesto la neutraliza. Comienzan a formarse micelas. La imagen en la pantalla del aula es entonces muy vívida, observándose un movimiento loco de gotas e hilos rojos hasta que la abundancia de micelas obstruye el paso de la luz a través del platillo de Petri.

Estas experiencias las finaliza Wissing con unas operaciones que son pura diversión. En una de ellas deposita en agua ácido oleico en cantidad suficiente para que persista una gota. Luego, en su centro, vierte una gota de Triton X-100, que es un detergente comercial. Cuando el detergente traspasa la gota de ácido y llega al agua, la superficie de la solución queda repentinamente limpia de ácido oleico.

En otro ensayo Wissing vierte sobre agua una diminuta gota de ácido oleico y agrega aceite de parafina suficiente para formar una gota de unos tres centímetros de diámetro. Acido oleico añadido adicionalmente junto al aceite de parafina rodea la gota, formando un complicado dibujo de gotas que prosiguen moviéndose durante mucho tiempo. En un experimento distinto, Wissing muestra de qué modo hacer maniobrar a una gota de ácido oleico en agua, valiéndose de un estropajo de algodón impregnado de hidróxido amónico concentrado que acerca a la gota. Dando al estropajo una sacudida rápida, la gota de ácido echa a correr a trompicones por la superficie del agua; y cuando la acorrala junto a la pared del platillo se pone a oscilar en el vapor que desprende el estropajo.

La formación de monocapas por los compuestos orgánicos parcialmente solubles desempeñó un papel primordial en las primeras investigaciones en torno al tamaño de las moléculas. Así, en 1890 Lord Rayleigh utilizó una cantidad muy pequeña de aceite de oliva para estimar el tamaño de sus moléculas. Mojó en el aceite un alambre fino de platino, que luego pesó en una balanza de precisión. Parte de ese aceite lo dejó libre sobre una superficie circular de agua de 84 centímetros de diámetro; volvió a pesar el alambre. La diferencia entre ambas pesadas constituía el peso del aceite que entonces cubría la superficie del agua.

Para determinar la extensión del desparramamiento de la monocapa roció el agua con alcanfor finamente molido. Como los gránulos de alcanfor se disuelven por uno de sus lados en el agua, rebajan la tensión superficial de ésta. Y la mayor tensión superficial existente en el otro lado hace que cada gránulo se mueva a tirones. Cuando añadía aceite de oliva, la tensión superficial resultaba demasiado débil para que los gránulos danzaran. Por tanto, podía así determinar la extensión de la capa de aceite observando los gránulos de alcanfor. Entonces, añadiendo la cantidad de aceite precisa para que los gránulos detuviesen su danza, conseguía cubrir de una monocapa uniforme la superficie de agua. Conociendo el área de la superficie y el peso de la gota calculó el espesor de la capa; unos 17 angstrom.

Rayleigh creyó, asimismo, que esa cifra representaba la longitud de las moléculas que formaban la monocapa.

Richard E. Crandall y Jean F. Delord, de Reed College, han desarrollado una modificación de la experiencia de Rayleigh para sus alumnos de física elemental. En este ejercicio el alumno puede calcular la longitud de la molécula de ácido oleico, la longitud de enlace media a lo largo de la molécula y la masa del átomo de carbono. Se empieza rociando una superficie de agua con polvo de licopodio. Cuando se agrega una gota de ácido oleico, la monocapa empuja el polvo hacia fuera, en forma circular; luego se contrae adoptando un perímetro irregular. Mas antes de que el círculo se contraiga, el alumno dibuja su contorno en papel cuadriculado. Este dibujo se mide entonces para saber el área del círculo.

Al alumno se le dan varios datos. La solución de ácido oleico se prepara con 200 partes por millón de ácido en alcohol, en volumen. (Cuando la gota se ha depositado, el alcohol se disuelve en el agua y no interviene en el desparramamiento del ácido.) El volumen de la gota es de unos 0,05 centímetros cúbicos. Al alinearse las moléculas de ácido con las cabezas hacia abajo y las colas hacia arriba, cada cabeza ocupa una superficie que puede estimarse en un cuadrado cuyo lado sea 0,1 la longitud de la molécula.

Las masas de O, C y H están en proporción 16:12:1. La densidad del ácido oleico es del orden de 0,895 gramos por centímetro cúbico.

Multiplicando la concentración del ácido oleico por el volumen de la gota, el alumno puede calcular el volumen de ácido que se esparce sobre el agua. Dividiendo este volumen por el área de la monocapa obtiene la altura de ésta, que es, asimismo, la longitud de la molécula de ácido.

Sigue un cálculo aproximado de la longitud de enlace, es decir, de la distancia entre dos átomos de carbono contiguos a lo largo de la molécula. Para simplificar, podemos considerar el enlace doble del centro de la molécula como si fuera simple: el resto de los enlaces a lo largo de la molécula son realmente simples. Podríamos también despreciar la estructura alineal de la cabeza de la molécula. De este modo trataremos con 17 enlaces entre átomos de carbono a lo largo de la molécula, o bien con 18 si incluimos el hidrógeno de la cola. Supongamos que los enlaces yacen en una recta. Cuando el número de ellos se reparte a lo largo de la molécula, encontramos que la longitud de enlace viene a resultar del orden de un angstrom.

Al multiplicar el volumen de ácido oleico por su densidad, se tiene la masa de la monocapa. El lector podrá averiguar la masa de cada molécula admitiendo que la cabeza de ésta sea un cuadrado de lado igual a 0,1 la longitud de la molécula. Como esta longitud ya la sabe, no le será difícil calcular el área cuadrada de la cabeza.

El número de moléculas de ácido oleico se calcula dividiendo el área de la monocapa por la superficie de una cabeza. Y dividiendo la masa de la monocapa por el número de moléculas tendremos la masa de una de éstas. Finalmente, la masa de un átomo de carbono aislado la obtendremos contando cuántos átomos de cada tipo existen en una molécula y empleando la relación conocida entre sus masas.

Por su pequeñez y complejidad, se hace difícil representar las moléculas. No obstante, gracias a las experiencias de Rayleigh sí podemos estimar su tamaño. El ejercicio de laboratorio de Crandall y Delord permite determinar la longitud de enlace. Con las experiencias de Wissing hasta podemos imaginar la ordenación de las moléculas en la superficie de separación entre el agua y un compuesto orgánico y entender por qué unos compuestos orgánicos se extienden sobre el agua en tanto que otros no.

Libros

Optica medieval, estadística bayesiana, formación de iones y genética y evolución de la cultura

Luis Alonso, Pilar García, Julio A. Alonso y Jesús Mosterín

Roger Bacon's Philosophy of Nature, por David C. Lindberg. Clarendon-Oxford University Press; Oxford, Londres, Nueva York, 1983. Reza, vertido al castellano, el subtítulo: edición crítica y traducción inglesa, con una introducción y notas de las obras *De multiplicatione specierum* y *De speculis comburentibus*.

La historiografía de la óptica medieval se divide en dos épocas: antes y después de Lindberg. O más exactamente, antes y después de *Theories of Vision from al-Kindī to Kepler* (Chicago, 1976), de dicho profesor de historia de la ciencia en la Universidad de Wisconsin. Distinguía allí una triple tradición: una orientación perspectivista o de óptica geométrica, una tradición aristotélica y una tradición teológica.

Constituían el núcleo de la primera Ibn al-Hayṯam (nuestro Alhazén), Roger Bacon, Witelo y John Pecham. De todos ellos se había venido ocupando Lindberg desde 1967. En algún caso había preparado su edición crítica (*Tractatus de perspectiva*, de Pecham; Nueva York, 1972); había hurgado incluso en temas centrales, como la refracción o los conos de luz. Y había escrutado las líneas maestras de la óptica medieval en el libro sobre teorías de la visión antes citado. Pero le faltaba, tras escaramuzas y escarceos, abordar el punto focal: luz y materia en Bacon.

Esa es la explicación del título del volumen. Núcleo de la filosofía de la naturaleza de Bacon es su doctrina de la causalidad natural, entendida como un proceso de radiación en el que la propagación de la luz constituía un ejemplo visible y, por ende, paradigmático. La introducción, fluida, se convierte en una clase magistral sobre cómo preparar una edición crítica: autoría de ambos libros, pluralidad y corrupción de los títulos, versiones manuscritas e impresas, fecha de composición y fuentes. Antes de llegar al examen de las obras escribe una sucinta biografía donde establece fechas, hasta ahora dudosas (nacimiento de Bacon en 1220, muerte en 1292 o poco después), acota estancias y viajes, depura relaciones (con Grosseteste) y amistades (con Guy de Foulques, el que sería al poco Clemente IV), manifiesta sus tendencias milenaristas y sus dificultades una vez ingresado en religión, si bien deja borroso el perfil humano del personaje: su talante demoledor (contra Alberto Magno, por ejemplo). Y fija la cronología del opus baconiano.

Viniendo ya al tema desarrollado en *De multiplicatione specierum* y *De speculis comburentibus* traza la historia "platónica" del mismo, deteniéndose en Filón, Plotino, los Padres de la Iglesia, el Pseudo-Dionisio y, sobre todo, al-Kindī, cuyo *De radiis* enseñó a los estudiosos medievales que toda creatura constituía una fuente de radiación, siendo el universo una inmensa red de fuerzas. No sería otra la doctrina central del *De multiplicatione specierum*.

El precedente inmediato de Bacon fue Robert Grosseteste (c. 1168-1253), para quien todo agente natural multiplica su potencia desde él mismo hasta el paciente que recibe, a su modo, dicha acción. Esa impronta es la especie. "No significará aquí especie, escribe Bacon, el quinto universal de Porfirio, sino que designará el primer efecto de cualquier agente que opere de una manera natural." Por decirlo con un ejemplo, especie de la luz que hay en el Sol es la luz solar en el aire, la claridad que atraviesa la ventana o los resquicios. La especie se asemeja a aquello que la produce. ¿Qué cosas producen especies? ¿Cómo se multiplican éstas, física y matemáticamente, a través de un medio? ¿Cómo se reciben las especies? ¿Cómo se debilitan y corrompen?

Dos son las propiedades principales de la especie: semejanza con el agente que la produce y univocidad de la misma. La semejanza de la especie a su agente se extiende hasta la naturaleza, definición y operación de éste, ya que el paciente –Bacon sigue a Aristóteles aquí– se halla en potencia ante el agente: el fuego produce fuego en algo (en el paciente, si usamos su expresión) que antes no ardía. Por univocidad se entiende la existencia de un solo primer efecto; efectos secundarios y terciarios habrá muchos: el primer efecto de la luz será siempre alumbrar, pero secundaria o subsidiariamente puede inducir calor, putrefacción y muerte.

Producen especies las cosas sensibles, la sustancia, los sentidos, los universales y los singulares. La producen por cambio de estado (no propiamente en sentido físico actual, aunque a veces pueda identificarse). La especie se multiplica a través de su generación a lo largo de las distintas partes del medio, generación en la que interviene la materia de dicho medio. Pero este medio (o paciente) tiene su propia naturaleza, que pretende conservar, y acabará por deshacerse de esa impronta, lo que significará la corrupción de la especie.

¿Qué relación guarda este galimatías de tufo escolástico con la óptica? Una muy estrecha. La multiplicación de la especie sigue un modelo de propagación de la luz. Modelo óptico que se miniaturiza en los espejos ustorios (*De speculis comburentibus*), según la *Catoptrica* de Euclides.

En las notas el lector hallará la explicación de voces técnicas cuyo conocimiento exacto se hace imperioso para no extrapolarlo a sus significados actuales: qualitas, simulacro, imago, lux, lumen, agens, patiens, renovatio, forma, radius incidens, radius conversus, tropica, etcétera. Además de las acotaciones históricas sobre el origen de las distintas ideas, las discordancias entre las diferentes versiones, etcétera.

Decíamos al comienzo, que *Theories of Vision* marcó un hito. Han comenzado las glosas y las matizaciones. Citaré dos: Peter Marshall ["Nicole Oresme in the Nature, Reflection, and Speed of Light", *Isis*, 72 (1981), 357-374] y Bruce Stansfield Eastwood ["The Elements of Vision: The Micro-Cosmology of Galenic Visual Theory According to Hunayn Ibn Isḥāq", *Transactions of the American Philosophical Society*, vol. 72, 1982]. Pero no aquietaría mi conciencia si dejara pasar por alto dos omisiones, una de ellas quizá comprensible –la del estudio de Millás Vallicrosa sobre Avicebrón– y otra, menos, los trabajos de Juan Vernet sobre el tema. En particular, el titulado "Dos tratados del Arquímedes árabe: *Tratado de los círculos tangentes* y el *Libro de los triángulos*" (Barcelona, 1979). Con su lectura se entiende mucho mejor la geometría óptica de los espejos en la tradición musulmana. (L. A.)

Bioestadística, una Perspectiva Bayesiana por J. M. Bernardo; Vicens-Vives, Barcelona, 1981. Ante la

necesidad de elegir entre la metodología bayesiana y la clásica para la formación estadística del futuro profesional, el autor se inclina hacia la primera, por estar exenta de contradicciones y por su unidad conceptual. En particular, considera las técnicas bayesianas relevantes para la práctica e investigación médica, merced a la gran información inicial con que se cuenta en la mayoría de los problemas médicos y ser éstos, con gran frecuencia, problemas de decisión.

El libro va dirigido al universitario que desee una formación moderna en metodología estadística y teoría de la decisión. Se trata de un curso introductorio que no exige conocimientos previos de estadística. Los párrafos de mayor dificultad matemática aparecen impresos en letra pequeña para su fácil identificación y pueden omitirse en una primera lectura sin perder el hilo. El libro se divide en siete capítulos. En el introductorio y primero detalla el alcance y objetivos, da unas primeras nociones sobre estadística, teoría de la decisión, probabilidad e inferencia. Defiende la metodología proporcionada por la teoría de la decisión frente a la estadística clásica, razonando sus principios fundamentales; justifica la introducción de la probabilidad desde un punto de vista bayesiano y advierte que no se ocupará de la estadística descriptiva, sino atenderá solamente a la inferencia.

Expone, en el siguiente, la probabilidad de un suceso A en una situación H, $p(A/H)$, como una medida del grado de creencia en A que nos sugiere la información contenida en H; incluye como caso particular las situaciones clásicas de existencia de simetrías y sucesos repetibles. Se introduce la utilidad como una función que mide las preferencias del decisor entre las posibles consecuencias y se razona que el modo natural de actuar es maximizando la utilidad esperada. Presenta unos principios de comportamiento coherente; si éstos son aceptados, dice, es necesario asignar una medida de probabilidad que describa la información del decisor y una función de utilidad que revele sus preferencias, eligiendo aquella decisión que maximice la utilidad esperada. Razona por qué se inclina por este principio de maximización de la utilidad y no por otros criterios de decisión.

Al abordar las probabilidades como grado de creencia, demuestra tres propiedades básicas que suelen darse como axioma en la definición de probabilidad clásica. Este hecho notable de que los grados de creencia se comporten de acuerdo con las mismas leyes que las frecuencias relativas funda una teoría de la probabilidad unificada.

Cuando se ocupa de la medida de probabilidad, tras introducir los conceptos de espacio probabilístico, independencia e intercambiabilidad, insiste en el modo de incorporar información mediante el teorema de Bayes y estudia algunos métodos para asignar probabilidades. Trata el capítulo cuarto de las cantidades aleatorias: variables discretas y continuas y principales distribuciones en ambos casos. Describe los conceptos de función de distribución, función de una cantidad aleatoria, momentos y función generatriz, dando una idea de su generalización al caso multivariante. Pasa por alto la función característica y los problemas de convergencia de sucesiones; sus definiciones de cantidad aleatoria y variable aleatoria continua quizá no acaben de satisfacer al profesional de la matemática.

A los conocimientos fundamentales expuestos sobre estos puntos por la mayoría de los libros se añaden aquí otros conceptos y resultados de interés: aproximación de la media y la varianza de una función de variable aleatoria y análisis secuencial obtenido al aplicar el teorema de Bayes.

En "el proceso de aprendizaje", expone la solución bayesiana a un problema de inferencia, describiendo el proceso que permite incorporar al análisis de un problema de decisión la información proporcionada por los datos experimentales relacionados con sus sucesos inciertos relevantes. Este proceso exige definir, mediante una distribución de probabilidad, la información inicial que se posee sobre los sucesos inciertos, precisar, mediante la función de verosimilitud, la relación entre los datos experimentales y estos sucesos, y determinar, mediante el teorema de Bayes, la distribución final que describe la información que se posee sobre las suceciones inciertas tras incorporar a la información inicial la que proporcionan los resultados experimentales.

Completa el estudio de la inferencia estadística, en el capítulo sexto, con una serie de aproximaciones que facilitan el análisis y la comunicación de los resultados y el estudio de las conclusiones que pueden extraerse de los datos cuando no se dispone de información inicial o no quiere utilizarse. En concreto se investigan las características y regiones creíbles que describen la distribución final; se dedica un apartado a las familias conjugadas, cuyo estudio permite un análisis rápido de los datos correspondientes a un modelo de la familia exponencial cuando es posible aproximar la información inicial por un elemento de la familia conjugada; se dan métodos para transformar la distribución final en otra aproximadamente normal; se sigue el comportamiento asintóticamente normal de la distribución final; se aportan, para el caso de ausencia de información inicial, las distribuciones finales de referencia y se comenta la estabilidad de la distribución final frente a cambios en los datos, el modelo o la distribución inicial.

Por último, en el capítulo 7 ("Análisis cuantitativo de decisiones") se aplican esas ideas a ciertos problemas de decisión. Desde un enfoque bayesiano acomete los problemas clásicos de estimación puntual y contraste de hipótesis y demuestra que la inferencia estadística puede considerarse un problema de decisión cuya función de utilidad sea una medida de información. Prosigue con los métodos para valorar la función de utilidad, determina el valor de la información proporcionada por los datos experimentales, aborda el problema del diseño de experimentos y estudia el tamaño muestral óptimo.

En resumen, el libro nos parece una valiosa perspectiva bayesiana de la estadística aplicada, de interés para profesionales de muy diversas áreas, y en particular para los médicos, que basen sus conclusiones científicas en el análisis de un conjunto de datos experimentales. No es, por tanto, el manual tradicional de bioestadística. Aun considerando que la metodología bayesiana ofrece una visión mucho más unificada y coherente que la clásica, nos parece arriesgado exponer sólo este enfoque, ya que por un lado pueden presentárseles muchos problemas cuya solución bayesiana hay que buscarla todavía en revistas especializadas o esté en estudio, y por otro lado la mayoría de los trabajos de la bibliografía médica que utilizan la estadística lo hacen desde un punto de vista clásico. (P. G. C.)

ION FORMATION FROM ORGANIC SOLIDS, por A. Benninhoven (compilador). Springer Verlag, Berlín, 1983. Este libro contiene las comunicaciones presentadas en la segunda conferencia internacional sobre formación de iones a partir de sólidos orgánicos, que tuvo lugar en la Universidad de Münster en septiembre de 1982. La conferencia se centró en la formación de iones a partir de sólidos orgánicos no volátiles. El progreso en este campo ha sido rápido, tras el descubrimiento de nuevos e inesperados procesos de ionización. Por constituir el primer libro que proporciona un tratamiento completo de las técnicas de ionización de materiales (especialmente orgánicos) no volátiles,

es de agradecer a Springer su celo en la publicación del mismo.

Se pretendía en la reunión adquirir, por un lado, un conocimiento básico de estos procesos de ionización débil y, por otro, examinar sus aplicaciones presentes y futuras. Y el campo de las aplicaciones es desde luego enorme. En el libro se describen numerosas aplicaciones a la física, química, bioquímica, medicina y farmacología. Seis artículos de revisión cubren las ideas fundamentales de los diversos procesos de ionización. Los artículos, todos ellos bastante completos, describen los dispositivos experimentales, las características de los espectros y la influencia de los parámetros de excitación; examinan los modelos propuestos para el proceso de formación de iones y evalúan las aplicaciones al análisis químico. El libro contiene veintiséis comunicaciones cortas sobre aspectos especializados.

El nexo común a todas las técnicas descritas es el siguiente: en un primer paso, y usando diversos procedimientos que comentaré brevemente, se desprenden de la muestra iones tales como M^+, $(M + H)^+$, $(M + Na)^+$, donde M designa cierto tipo de molécula existente en la muestra. También se forman iones negativos que proporcionan información complementaria. En un segundo paso, estos iones moleculares pasan a un espectrómetro de masas, donde el espectro obtenido servirá para identificar la molécula. La principal aplicación ha sido la determinación de pesos moleculares (los espectros identifican iones moleculares con una relación masa/carga eléctrica de hasta 5000 e incluso mayor) y en menor escala la identificación de estructuras moleculares. Las diversas técnicas se refieren a las formas de producir esos iones.

La primera parte se ocupa de la técnica conocida por espectrometría de masas por desorpción inducida por un campo (FDMS). La describen F. W. Röllgen y H. R. Shulten, de la Universidad de Bonn. En esta técnica, que data del año 1966, la muestra se deposita en disolución sobre la superficie de un ánodo. La ionización y el desprendimiento de los iones moleculares se debe al efecto combinado de un alto voltaje y del calentamiento del ánodo. Röllgen expone los detalles experimentales y analiza los mecanismos de formación de los iones moleculares. En este capítulo, queda patente que el mecanismo de formación de iones moleculares es un tema abierto y controvertido en el que el lector interesado en la química teórica encontrará numerosas sugerencias. Shulten estudia las aplicaciones analíticas de la FDMS en física, química, biología y medicina.

Una fracción substancial del libro se dedica a la espectrometría de masas de iones secundarios (SIMS). Técnica que consiste en bombardear la muestra con iones acelerados (es decir, iones con gran energía cinética). La energía se disipa por colisiones a lo largo de la trayectoria del ion. Una cierta porción de esta energía puede retornar a la superficie de la muestra y resultar en el desprendimiento de una partícula (molécula o fragmento molecular). Parte de estas partículas están cargadas, lo que constituyen la base del SIMS. A. Benninghoven, de Münster, pasa revista a los mecanismos de interacción del ion incidente con la zona superficial de la muestra. R. D. McFarlane, de la Universidad de Texas A & M, iniciador de la SIMS en 1974, pondera los resultados obtenidos al usar como proyectiles los fragmentos de fisión del ^{252}Cf. El bombardeo con átomos rápidos es presentado por C. Fenselau, de la Universidad Johns Hopkins. El dispositivo experimental empleado en esta técnica es más flexible y sencillo que los correspondientes a otras formas de SIMS. Aunque data de fecha recentísima, 1981, existen ya fuentes comerciales que producen átomos neutros con energías del orden de 10 Kv. Xenon suele ser el elemento preferido. Los trabajos específicos que describen las aplicaciones de la SIMS abarcan el estudio de diversos materiales: metales, polímeros, proteínas y moléculas orgánicas. F. Hillenkamp, de la Universidad de Frankfurt, expone la formación de iones inducida por láser y su aplicación al estudio de moléculas de interés biológico.

La última parte de la obra se refiere a la emisión de iones por líquidos. Lleva la firma de M. L. Vestal, de la Universidad de Houston. Esta técnica consiste en disolver la muestra en un disolvente líquido y someter después la superficie del mismo a un campo electromagnético intenso, lo que origina el desprendimiento de partículas ionizadas, que pasan entonces al espectrómetro de masas. A veces, la disolución es primero pulverizada y posteriormente sometida al campo eléctrico.

La novedad de las técnicas de análisis y la variedad de aplicaciones ofrecida convierten este volumen en un libro imprescindible para todo grupo de investigación interesado en el análisis químico de materiales orgánicos y en la física o química de superficies. La mayoría de los artículos son asequibles a un lector con una buena formación en química, o al menos familiarizado con la química física o con la física de sólidos. Las técnicas de análisis tratadas usan una tecnología bastante refinada, que sería deseable empezar a encontrar en nuestras universidades y laboratorios si queremos alcanzar el nivel que todos deseamos. Encuentro recomendable la lectura de varios artículos para químicos teóricos, que pueden hallar problemas en los que aplicar métodos de la química cuántica al estudio de estos materiales complejos, sobre todo en lo que se refiere a los mecanismos de excitación, fragmentación e ionización de las moléculas de la muestra. (J. A. A.)

La Evolución de la Cultura en los Animales, por John T. Bonner. Alianza Editorial, Madrid, 1982. El desarrollo de la cosmovisión científica del mundo, desde Copérnico hasta nuestros días, ha estado caracterizada por la creciente decadencia del inicial antropocentrismo. La noticia de que la Tierra no estaba en el centro del universo (como correspondía a la dignidad de su principal habitante, el hombre), sino que era un planeta más que giraba alrededor del Sol, produjo en su tiempo un verdadero *shock* emocional. Los posteriores descubrimientos de que el Sol no era tampoco sino una estrella periférica cualquiera de nuestra galaxia, y que nuestra galaxia no era sino otra más entre los miles de millones de galaxias, acabaron de arrinconar definitivamente cualquier pretensión antropocéntrica en astronomía. Irremediablemente expulsado de la cosmología, el antropocentrismo se refugió en la biología. Pero la revolución darwiniana supuso el reducir la especie humana a una rama más del gran árbol de la vida, rama estrechamente emparentada con las demás y surgida mediante los mismos procesos y presiones ambientales que las otras. El antropocentismo tuvo que ceder posiciones y retirarse a la psicología y la sociología, que cubrirían fenómenos únicos de la especie humana e incomparables con el resto del mundo animal. El desarrollo primero de la etología y luego de la sociobiología ha dado también al traste con esas pretensiones. Muchos de los resortes básicos de nuestra conducta individual y social los compartimos con otras especies. Y sin embargo es evidente que actualmente los humanos ocupamos una posición preponderante en este planeta, posición que debemos al extraordinario desarrollo de nuestra cultura.

John T. Bonner, conocido biólogo de la Universidad de Princeton, dedica el libro aquí reseñado a dibujar un cuadro

general de la cultura, sus condiciones y sus orígenes, en el que quepa tanto la nuestra como la de los otros animales.

El libro se divide en 8 capítulos. El primero presenta la definición de cultura como transferencia de información por medios conductales, sobre todo mediante el aprendizaje. El capítulo segundo elabora muy elegante y precisamente la diferencia y contraste entre evolución genética y evolución cultural. La evolución genética se presenta en términos de los genes "egoístas", según la metáfora de Dawkins, como la sucesión de las máquinas de supervivencia (los organismos) que se construyen los genes para sobrevivir. En la evolución genética la información es transmitida por los genes; en la evolución cultural, por los "memes" (otro término de Dawkins), que son las porciones de información pasadas conductalmente de un individuo a otro. Aunque ambas evoluciones están guiadas por la selección natural, las diferencias son obvias. En la evolución genética un individuo transmite información sólo a sus descendientes y sólo en el momento de su fecundación; en la evolución cultural, un individuo puede transmitir o recibir información de muchos otros. Los cambios en la evolución genética son muy lentos; en la cultural pueden ser muy rápidos.

Los dos sistemas físicos responsables de la evolución genética y la cultural son, respectivamente, el genoma y el cerebro, y de ellos trata el capítulo tercero. Ambos son sistemas capaces de recibir, almacenar, procesar y transmitir información. El cerebro es un invento del genoma, al que sirve para lograr una adaptación rápida y fácil del organismo a las cambiantes condiciones del medio. Bonner estudia la evolución del sistema nervioso, y en especial del cerebro humano, que explica por un fenómeno de neotenia, es decir, de retención de un carácter juvenil. Hay muy pocas diferencias entre los genes de un chimpancé y los de un humano. El cambio genético que condujo a nuestro gran cerebro debió afectar a muy pocos genes. En efecto, no fue necesario diseñar un nuevo cerebro, bastó con alargar el período de rápido crecimiento juvenil del cerebro, que también se da en los otros antropoides, pero que en ellos se interrumpe mucho antes que en nosotros.

Los capítulos cuarto a séptimo están dedicados a analizar los diversos pasos evolutivos que condujeron a la cultura, y a través de los cuales la inicial rigidez fue siendo sustituida por una cada vez mayor flexibilidad en la adaptación al medio. Incluso la evolución genética tuvo que flexibilizarse al principio, pasando de la rígida reproducción vegetativa a la reproducción asexual con mutaciones y, sobre todo, a la reproducción sexual con mutaciones y recombinación, que fue la clave de todos los progresos posteriores. El capítulo cuarto considera los orígenes de la evolución cultural, cuyos remotos inicios aparecen en las bacterias móviles, capaces ya de recibir información del entorno y reaccionar en consecuencia. El capítulo quinto estudia la evolución de las sociedades animales, desde los enjambres bacterianos hasta las sociedades de insectos y de mamíferos, y trata también de la evolución de los sistemas de comunicación. El capítulo sexto está dedicado a la evolución del aprendizaje y la enseñanza por ensayo y error, la enseñanza por el ejemplo y la enseñanza por el lenguaje. En el capítulo séptimo se estudia la progresiva flexibilización de la conducta desde las pautas de comportamiento de respuesta única (las acciones reflejas) hasta las pautas de comportamiento de elección múltiple entre varias respuestas, que implican una decisión entre ellas por parte del cerebro, e incluso hasta la invención de respuestas nuevas, que es lo que posibilita la cultura.

Después de haber rastreado la evolución de sus condicionantes biológicos, el capítulo octavo y último trata directamente de la evolución de la cultura misma. Aquí aparece el uso de herramientas (de espinas por los pinzones de las Galápagos para sacar larvas de los árboles; de piedras por las focas para romper conchas de moluscos; de ramitas preparadas por los chimpancés para sacar termitas, etc.) y de técnicas como las empleadas por los ostreros para abrir un mejillón, la transmisión de información sobre predadores (como los elefantes africanos mansos que, después de verse cazados por el hombre, se vuelven esquivos y agresivos), el aprendizaje de dialectos (en diversas aves) y el conocimiento de rutas y lugares. Especialmente interesantes son los casos de invención y transmisión por imitación de nuevas prácticas, como cuando a un herrerillo se le ocurrió agujerear con el pico la tapa de una de las botellas de leche que dejan a la puerta de las casas en Inglaterra y comerse la nata, costumbre luego progresivamente adoptada por sus congéneres hasta convertirse en una plaga; o el bien conocido caso de los inventos de la macaca japonesa Imo, como el de la limpieza y salado de las batatas, y luego adoptados por todos los macacos de su isla. También considera la cultura humana, enormemente potenciada por su lenguaje flexibilísimo y por métodos de almacenamiento acumulativo como la escritura. Finalmente vuelve al tema constante del libro: la explicación evolucionaria de la cultura. La presión selectiva de un entorno cambiante sobre los genes es la que ha conducido a cerebros cada vez mayores y mejores, a sistemas de sociabilidad, comunicación y cultura que permiten respuestas más rápidas y eficaces.

El libro de Bonner está bien escrito, bien argumentado y lleno de casos y detalles iluminadores. Da gusto leerlo. Aunque explícitamente rehuye los temas más polémicos en que se intersectan sociología y sociobiología, mantiene constantemente un vigoroso enfoque evolucionario y darwinista, claramente influido por R. Dawkins y E. O. Wilson en sus aspectos más teóricos. Desde un punto de vista conceptual, sólo le criticaría el que defina (al principio del capítulo quinto) tanto la cultura como la comunicación del mismo modo, como transferencia o transmisión de información por medios conductales, aunque no las identifique, pues dice, por ejemplo, que la cultura depende de la comunicación; y el que no defina en absoluto la información. Pero quizás exigir aquí definiciones precisas de términos tan difíciles como información y comunicación sería pedir peras al olmo.

La traducción española del libro, debida a Natividad Sánchez, no es mala ni buena. Contiene algunos olvidos, como el de una oración en la página 24, y algunas traducciones falsas, como la de *flatworm* (platelminto) por "lombriz" (p. 23), *quibble* (sutileza) por "sofisma" (p. 89) o *muted* (apagado) por "conmutaciones" (p. 135). Un punto positivo es que ha añadido el nombre latino de las especies citadas que carecen de nombre usual en castellano. La edición del libro no es muy cuidada. En la página 74 se ha variado la disposición de los dibujos, colocándolos todos uno encima de otro, pero el pie sigue remitiendo a la derecha y la izquierda. Más grave es lo sucedido en la página 168, donde se ha dado la vuelta al grabado, conservando el mismo pie, con lo que el ave que está a la izquierda ya no es el macho, como dice el pie, sino la hembra. El índice general de la edición original ha desaparecido en la española. Afortunadamente se han conservado los magníficos dibujos puntillados de Margaret La Farge, aunque desgraciadamente han resultado un tanto ennegrecidos en la impresión. (J. M.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

LARVAS DE CEFALOPODOS

A REVIEW OF THE SYSTEMATICS AND ECOLOGY OF OCEANIC SQUIDS. M. R. Clarke en *Advances in Marine Biology*, vol. 4, págs. 91-300; 1966.

SYSTEMATICS, DISTRIBUTION, AND ABUNDANCE OF THE EPIPLANKTONIC SQUID (CEPHALOPODA, DECAPODA) LARVAE OF THE CALIFORNIA CURRENT. April 1954 - March, 1957. T. Okutani y J. McGowan en *Bulletin of the Scripps Institution of Oceanography*, vol. 14; 1969.

CEPHALOPODS AND FISH: THE LIMITS OF CONVERGENCE. A. Packard en *Biological Review*, n.º 47, págs. 241-307; 1972.

THE LARVAE OF CEPHALOPODA: A REVIEW. S. von Boletzky en *Thalassia Jugoslavica*, volumen 10 (1/2), páginas 45-76; 1974.

THE BIOLOGY OF CEPHALOPODS. SYMPOSIA OF THE ZOOLOGICAL SOCIETY OF LONDON, n.º 38. Dirigido por M. Nixon y J. B. Messenger. Academic Press, 1977.

IO

THE ROLE OF SO_2 IN VOLCANISM ON IO. B. A. Smith, E. M. Shoemaker, S. W. Kieffer y A. F. Cook II en *Nature*, vol. 280, n.º 5725, págs. 738-743; 30 de agosto de 1979.

THE GALILEAN SATELLITES AND JUPITER: VOYAGER 2 IMAGING SCIENCE RESULTS. Bradford A. Smith, Laurence A. Soderblom et al. en *Science*, vol. 206, n.º 4421, págs. 927-950; 23 de noviembre de 1979.

LOS SATÉLITES GALILEANOS DE JÚPITER. Laurence A. Soderblom en *Investigación y Ciencia*, n.º 42, págs. 58-71; marzo, 1980.

HEAT FLOW FROM IO (JI). Dennis L. Matson, Gary A. Ransford y Torrence V. Johnson en *Journal of Geophysical Research*, vol. 86, n.º B3, págs. 1664-1672; 10 de marzo de 1981.

SATELLITES OF JUPITER. Dirigido por D. Morrison. University of Arizona Press, 1982.

TWO CLASSES OF VOLCANIC PLUMES ON IO. Alfred S. McEwen y Laurence A. Soderblom en *Icarus*, vol. 55, n.º 2, págs. 191-217; agosto, 1983.

UNA COMUNIDAD AGRICOLA DE LA EDAD DEL HIERRO EN EUROPA CENTRAL

EUROPEAN PREHISTORY. Sarunas Milisauskas. Academic Press, 1978.

FROM BRONZE AGE TO IRON AGE: ECONOMIC, HISTORICAL AND SOCIAL CONSIDERATIONS. R. Peroni en *Italy before the Romans: The Iron Age, Orientalizing, and Etruscan Periods*, dirigido por David y Francesca R. Ridgway. Academic Press, 1979.

RURAL ECONOMY IN THE EARLY IRON AGE: EXCAVATIONS AT HASCHERKELLER 1978-1981. Peter S. Wells. Museo Peabody de la Universidad de Harvard, 1983.

LA HELICE DEL ADN Y SU LECTURA

MECHANICS OF SEQUENCE-DEPENDENT STACKING OF BASES IN *B*-DNA. C. R. Calladine en *Journal of Molecular Biology*, vol. 161, n.º 2, págs. 343-352; 25 de octubre de 1982.

ORDERED WATER STRUCTURE AROUND A *B*-DNA DODECAMER: A QUANTITATIVE STUDY. Mary L. Kopka, Albert V. Fratini, Horace R. Drew y Richard E. Dickerson en *Journal of Molecular Biology*, vol. 163, n.º 1, págs. 129-146; 5 de enero de 1983.

SEQUENCE-DEPENDENT CONFORMATION OF AN *A*-DNA DOUBLE HELIX. Z. Shakked, D. Rabinovich, O. Kennard, W. B. T. Cruse, S. A. Salisbury y M. A. Viswamitra en *Journal of Molecular Biology*, vol. 166, n.º 2, págs. 183-201; 15 de mayo de 1983.

BASE SEQUENCE AND HELIX STRUCTURE VARIATION IN *B* AND *A* DNA. Richard E. Dickerson en *Journal of Molecular Biology*, vol. 166, número 3, páginas 419-441; 25 de mayo de 1983.

GRAVEDAD CUANTICA

GENERAL RELATIVITY: AN EINSTEIN CENTENARY SURVEY. Dirigido por S. W. Hawking y W. Israel. Cambridge University Press, 1979.

QUANTUM GRAVITY II. Dirigido por C. J. Isham, R. Penrose y D. W. Sciama. Oxford University Press, 1981.

ALTERNANCIA DE PRESAS EN UN ECOSISTEMA SIMPLE

THE DISTRIBUTION AND ABUNDANCE OF ANIMALS. H. G. Andrewartha y L. C. Birch. University of Chicago Press, 1954.

ECOLOGICAL GENETICS. E. B. Ford, John Wiley & Sons, Inc., 1965.

OF PREDATION AND LIFE. Paul L. Errington. Iowa State University Press, 1967.

AERODINAMICA DE LOS VEHICULOS TERRESTRES DE PROPULSION HUMANA

PROCEEDINGS OF THE FIRST HUMAN POWERED VEHICLE SCIENTIFIC SYMPOSIUM. Dirigido por Alan Abbott. IHPVA, P.O. Box 2068, Seal Beach, Calif. 90740; 1981.

BICYCLES & TRICYCLES: AN ELEMENTARY TREATISE ON THEIR DESIGN AND CONSTRUCTION. Archibald Sharp. Reimpresión de la edición de 1896 por MIT Press, 1982.

BICYCLING SCIENCE, Frank R. Whitt y David G. Wilson. The MIT Press, 1982.

INTERPRETACION DE LAS ILUSIONES OPTICAS

DIFFERENTIAL GEOMETRY OF CURVES AND SURFACES. Manfredo P. do Carmo. Prentice-Hall, Inc., 1976.

VISION: A COMPUTATIONAL INVESTIGATION INTO THE HUMAN REPRESENTATION AND PROCESSING OF VISUAL INFORMATION. David Marr. W. H. Freeman and Company, 1982.

THE MODULARITY OF MIND. Jerry Fodor. The MIT Press, 1983.

JUEGOS DE ORDENADOR

MATHEMATICAL STRUCTURES COMPUTER SCIENCE. Judith L. Gersting. W. H. Freeman and Company, 1967.

ALGORITHMS + DATA STRUCTURES = PROGRAMS. Niklaus Wirth. Prentice-Hall, Inc., 1976.

INTRODUCTION TO AUTOMATA THEORY, LANGUAGES, AND COMPUTATION. John E. Hopcroft y Jeffrey D. Ullman. Addison-Wesley Publishing Company, 1979.

TALLER Y LABORATORIO

INTERATOMIC AND INTERMOLECULAR FORCES. John P. McTague en *Encyclopedia of Physics*, dirigida por Rita G. Lerner y George L. Trigg. Addison-Wesley Publishing Company, Inc., 1981.